

EL ESCENARIO DE LA QUÍMICA CUÁNTICA*

Juan Carlos Paniagua

Departament de Química Física. Universitat de Barcelona

Correo electrónico: jpaniagua@ub.edu

Barcelona, 1990 (revisado el 18 de octubre de 2007)

PROLOGO

Existen textos excelentes sobre química cuántica que abarcan desde el nivel introductorio, orientado al que se enfrenta por primera vez con la materia, hasta niveles avanzados adecuados para una puesta al día del investigador especializado. No obstante, resulta a veces difícil conectar las herramientas que se utilizan para desarrollar la química cuántica en dichos textos con los principios en los que se basa esta disciplina, tal como se enuncian en la mayor parte de los libros de mecánica cuántica orientados a estudiantes de física. Por ejemplo, el formalismo de matrices de densidad reducidas desarrollado, entre otros autores, por Löwdin y McWeeny está estrechamente emparentado con los operadores de densidad que se utilizan para describir estados mezcla en mecánica estadística cuántica, pero la forma en que suelen introducirse uno y otros no evidencia fácilmente la conexión que existe entre ellos. Esto puede conducir a interpretaciones erróneas, como ocurre con la incorrectamente denominada matriz de densidad de primer orden en una base no ortogonal de orbitales (véase §3.5). Por otra parte, la descripción de estados mezcla mediante operadores de densidad ha cobrado renovado interés como consecuencia de la revolución provocada en la resonancia magnética nuclear por las técnicas multi-impulsionales. En efecto, la explicación de muchas de estas técnicas exige la utilización de aquel formalismo, que no suele incluirse ni en los currícula de las licenciaturas en química ni en los textos de química cuántica de nivel de post-grado.

Las consideraciones anteriores nos han llevado a alabar estos apuntes de introducción a aquella disciplina en los que se incide especialmente en la conexión de las herramientas que necesita el químico cuántico con los principios en los que se fundamenta la mecánica cuántica. Muchos usuarios de la química cuántica opinarán que no se precisa un entendimiento detallado

*Este texto está sujeto a una licencia de Reconocimiento-NoComercial-CompartirIgual (by-nc-sa) Creative Commons. Para más información véase <http://creativecommons.org/licenses/by-nc-sa/3.0/>

de las bases del formalismo para poder aplicarlo, y probablemente estén en lo cierto, pero es indudable que una formación básica sólida proporciona un sentido crítico sumamente valioso para analizar cualquier resultado que se aparte de los parámetros convencionales.

El planteamiento de estas notas es claramente introductorio, por lo que no pretendemos desarrollar en profundidad cada concepto ni detallar las aplicaciones prácticas. Tampoco pretendemos cubrir de manera exhaustiva la metodología de la química cuántica. Por ejemplo no mencionaremos los métodos perturbativos estacionarios ni dependientes del tiempo ni los problemas de colisiones, aplicaciones de la teoría de gran importancia en la química que se encuentran excelentemente desarrolladas en diversos textos. Tampoco utilizaremos el formalismo de segunda cuantización que, si bien es una herramienta muy práctica para desarrollar ciertos aspectos de la química cuántica, no introduce elementos físicos nuevos y puede enmarcarse en la teoría de representaciones que será desarrollada con cierto detalle.

En ocasiones hemos sacrificado el rigor matemático en favor de la sencillez sin que ello condicione la aplicabilidad del formalismo en el contexto en el que va a ser utilizado. Por ejemplo, hemos omitido o relegado a anotaciones a pie de página ciertas precisiones referentes a los dominios de definición de los operadores (por ejemplo, la distinción entre operadores autoadjuntos y hermíticos), a la existencia de reglas de superselección que puedan limitar la validez de un enunciado y a la extensión al continuo de resultados obtenidos para operadores de espectro discreto. Asimismo, hemos discutido la simetrización de los vectores de estado de sistemas que contienen partículas indistinguibles sin entrar en la consideración de las paraestadísticas y hemos introducido el colapso del estado sin hacer referencia a las recientes formulaciones de la mecánica cuántica "sin reducción", las cuales aportan interesante información sobre uno de los aspectos más espinosos de la formulación estándar.

Por otra parte, ciertos teoremas y propiedades se enuncian sin demostración y otros han sido relegados a ejercicios intercalados en el texto. Aunque no es preciso resolver todos los problemas para poder seguir del hilo argumental, es muy recomendable leer siempre los enunciados, ya que, al margen de su valor ilustrativo, pueden contener resultados que se utilicen posteriormente en el texto. La inclusión de complementos matemáticos intercalados en el texto, en lugar de dedicarles un capítulo introductorio o un apéndice, crea ciertos problemas en relación con su ordenación lógica, pero facilita su asimilación tanto por la forma dosificada en la que se introducen como por la inmediatez de su aplicación.

Aunque utilizaremos la notación de Dirac para los productos escalares, no introduciremos el "bra" correspondiente a un "ket" como elemento del espacio dual dado que, si bien es práctica habitual en los textos de mecánica cuántica, lo consideramos una complicación innecesaria para el desarrollo de la materia que trataremos.

Dada la comodidad que ello supone en el desarrollo de la química cuántica, utilizaremos con frecuencia unidades atómicas, las cuales pueden obtenerse a partir del sistema internacional

definiendo tres magnitudes mecánicas y una eléctrica (las unidades de temperatura, cantidad de materia e intensidad luminosa no se modifican), por ejemplo,

unidad atómica de masa = masa en reposo del electrón : $m_e = 9,109\ 389\ 7 \times 10^{-31}$ kg

unidad atómica de momento angular = constante de Planck/ 2π : $\hbar = 1,054\ 572\ 66 \times 10^{-19}$ J s

unidad atómica de longitud = bohr : $a_0 = 5,291\ 772\ 49 \times 10^{-11}$ m

unidad atómica de carga = carga elemental : $e = 1,602\ 177\ 33 \times 10^{-19}$ C .

El nivel del texto se adecúa a los conocimientos de un licenciado en química o física; en particular, se suponen conocimientos elementales de matemáticas al nivel de un primer ciclo de cualquier carrera científico-técnica (nociones de espacio vectorial, cálculo matricial elemental, etc.). La comprensión del texto resultará más fácil si se dispone de alguna formación previa en mecánica cuántica, al nivel de las asignaturas troncales de las licenciaturas en química o física, o de los textos de introducción a la química cuántica. No obstante, dicha formación no es un requisito indispensable, ya que la exposición de la teoría arrancará desde cero. Ocasionalmente se utilizará, sin demostración, algún resultado de teoría de grupos.

Dado el carácter general de los temas tratados en el texto, no se han incluido referencias bibliográficas salvo en casos en los que se hace alusión directa al contenido de alguna publicación.

INDICE

1. POSTULADOS DE LA MECANICA CUANTICA	1
1.0 Introducción	1
1.1 Primer postulado: Correspondencias sistema \leftrightarrow espacio de Hilbert y estado puro \leftrightarrow rayo unitario	2
1.2 Segundo postulado: Correspondencia observable \leftrightarrow operador	5
1.3 Tercer postulado: Resultados de medidas	7
1.4 Cuarto postulado: Colapso del estado	14
1.5 Quinto postulado: Evolución libre	18
1.6 Sexto postulado: Sistemas compuestos	22
1.6.1 Sistemas N-electrónicos	28
2. REPRESENTACIONES	31
2.0 Introducción	31
2.1 Bases discretas	31
2.1.1 Cambios de base	33
2.1.2 Valores y vectores propios	34
2.1.3 Ortonormalización	37
2.2 Bases truncadas	38
2.3 Bases continuas	41
2.4 Notación	44
3. MATRICES DE DENSIDAD REDUCIDAS	47
3.0 Introducción	47
3.1 Operadores de densidad reducidos	47
3.2 Matrices de densidad reducidas de sistemas polielectrónicos en representación de posiciones	49
3.2.1 Matriz de densidad reducida de orden 1	49
3.2.2 Matriz de densidad reducida de orden 2	51
3.3 Matrices de densidad reducidas de sistemas polielectrónicos en bases discretas	54
3.3.1 Funciones de onda multideterminantales	55
3.3.2 Funciones de onda monodeterminantales	57
3.4 Spin-orbitales naturales	59
3.5 Análisis poblacional	63

1. POSTULADOS DE LA MECANICA CUANTICA

1.0 Introducción

Entre los diversos aspectos en los que la *mecánica cuántica* choca con las ideas clásicas, cabe destacar los relacionados con la medición de observables.

En las teorías precuánticas se admite la posibilidad de medir con precisión, en principio, ilimitada variables suficientes para que cualquier observable quede determinado en instantes posteriores, siempre que se conozcan las interacciones que determinan la evolución del sistema. Por ejemplo, para un sistema mecánico aislado bastará conocer en un instante determinado las posiciones y los momentos de las partículas que lo constituyen y las fuerzas que actúan entre éstas. También se supone que los procedimientos de medición siempre pueden refinarse de modo que no alteren de modo significativo el estado del sistema.

De acuerdo con la teoría cuántica, en cambio, no es posible efectuar mediciones que determinen unívocamente todas las propiedades de un sistema ni su evolución futura, a pesar de que –contrariamente a una concepción bastante extendida–, no se establece ningún límite en la precisión con la que puede medirse cualquier observable. Deberemos, pues, conformarnos con conocer distribuciones de probabilidad de los valores que podrán tomar los observables. Este carácter probabilístico de las predicciones es uno de los factores que hacen necesario un modelo matemático para desarrollar la teoría fundamentalmente distinto del utilizado en mecánica clásica. Por otra parte, los efectos de las mediciones sobre el estado del sistema observado no son, en general, irrelevantes; al contrario, juegan un papel crucial en algunas de las aplicaciones más prometedoras de la teoría

En los años 1925-26 se propusieron, de forma independiente y casi simultánea, dos formulaciones de la moderna teoría cuántica con enfoques muy diferentes: la *mecánica de matrices* de Heisenberg y la *mecánica ondulatoria* de Schrödinger, cuya equivalencia fue probada por Schrödinger en 1926. En 1927 von Neumann y Landau (según otras fuentes también Bloch) extendieron la teoría a sistemas parcialmente determinados. Dirac desarrolló, ese mismo año, una teoría cuántica del campo electromagnético, y en 1928 incorporó las ideas relativistas a la mecánica cuántica, iniciando así el desarrollo de un marco teórico más general que desembocaría, más tarde, en la *electrodinámica cuántica*. Esta teoría describe completamente –por lo que hasta ahora se conoce– las interacciones involucradas en los procesos químicos.

En 1930 Dirac publicó un libro –*The Principles of Quantum Mechanics*– en el que se establece una formulación de la mecánica cuántica –podríamos denominarla *mecánica de vectores*– y una notación que serían rápidamente adoptadas por la comunidad científica. Las

formulaciones de Heisenberg y de Schrödinger se obtienen de forma muy directa a partir de la Dirac, la cual fue rigORIZADA por von Neumann en un texto definitivo –*Mathematische Grundlagen der Quantenmechanik*– publicado en 1932. En él se sentaron también las bases de la teoría de la medida, uno de los aspectos más controvertidos de la mecánica cuántica.

En el presente capítulo expondremos las bases de la mecánica cuántica siguiendo una línea argumental basada en el texto de von Neumann y utilizando una notación similar a la propuesta por Dirac. Nos limitaremos a desarrollar la formulación no relativista de la teoría, dado que los efectos relativistas no juegan un papel esencial en la mayor parte de las aplicaciones de la teoría a la química y, cuando ello es necesario, pueden incorporarse a posteriori en un tratamiento no relativista. Desarrollaremos la teoría sobre la base de seis postulados, de los cuales los dos primeros definen el marco matemático en el que se representan los estados y las propiedades de un sistema físico, el tercero indica como extraer información práctica a partir de los elementos representativos previamente introducidos, el cuarto y el quinto versan sobre la evolución temporal del sistema y el sexto establece importantes matizaciones relativas a sistemas constituidos por subsistemas diferenciables.

1.1 Primer postulado: Correspondencias sistema ↔ espacio de Hilbert y estado puro ↔ rayo unitario

A cada sistema físico se le hace corresponder un espacio de Hilbert complejo y separable (\mathcal{H}). A cada estado puro del sistema en un instante (t) se le hace corresponder un rayo unitario perteneciente al espacio de Hilbert asociado. Cualquier elemento de dicho rayo (Ψ_t) puede utilizarse para representar o describir el estado puro y se le da el nombre de "vector de estado" o "ket".

Comentarios:

i) Sin entrar en precisiones matemáticas, podemos definir un espacio de Hilbert complejo como un espacio vectorial definido sobre el cuerpo escalar \mathcal{C} de los números complejos y dotado de un producto interno o escalar que, para dos vectores ψ y ψ' de \mathcal{H} , designaremos mediante la notación $\langle \psi | \psi' \rangle$:

$$\langle \psi | \psi' + \psi'' \rangle = \langle \psi | \psi' \rangle + \langle \psi | \psi'' \rangle \quad \forall \psi, \psi', \psi'' \in \mathcal{H},$$

$$\langle \psi | k \psi' \rangle = k \langle \psi | \psi' \rangle \quad \forall \psi, \psi' \in \mathcal{H}, \quad \forall k \in \mathcal{C},$$

$$\langle \psi | \psi' \rangle = \langle \psi' | \psi \rangle^* \quad \forall \psi, \psi' \in \mathcal{H},$$

$$\langle \psi | \psi \rangle \geq 0 \quad \forall \psi \in \mathcal{H},$$

$$\langle \psi | \psi \rangle = 0 \quad \Leftrightarrow \quad \psi = 0.$$

El calificativo *separable* significa que contiene un conjunto numerable *denso*, es decir, un conjunto discreto de vectores $\{\psi_1, \psi_2, \dots\}$ que permiten aproximar cualquier vector ψ del espacio tanto como queramos mediante combinaciones lineales de los elementos del conjunto; tal aproximación debe entenderse con la noción de distancia que se deriva del producto escalar:

$$\lim_{n \rightarrow \infty} \left\langle \psi - \sum_{i=1}^n c_i \psi_i \mid \psi - \sum_{i=1}^n c_i \psi_i \right\rangle = 0,$$

lo cual indicaremos abreviadamente así:

$$\psi = \sum_{i=1}^{\infty} c_i \psi_i.$$

Diremos que un conjunto $\{\psi_1, \psi_2, \dots\}$ con dicha propiedad es *completo* o que es una *base* del espacio \mathcal{H} . Con frecuencia nos interesará expresar los vectores del espacio en función de conjuntos que, si bien no son completos, pueden "completarse" añadiendo un conjunto continuo de "vectores" ψ_x que no pertenecen al espacio (representan situaciones límite no realizables físicamente), y hablaremos, impropriadamente, de "bases" continuas o con parte continua:

$$\psi = \sum_i c_i \psi_i + \int_x \psi_x c(x) dx.$$

ii) Un rayo unitario es un subconjunto de \mathcal{H} cuyos elementos se pueden expresar de la forma $(e^{i\alpha} \psi)_{\alpha \in \mathbb{R}}$, siendo ψ un vector normalizado fijo de \mathcal{H} ($\langle \psi \mid \psi \rangle = 1$) y α un número real arbitrario.

iii) Para entender el significado de "estado puro" es conveniente considerar la forma en que se prepara un sistema que va a ser objeto de estudio. Dicha preparación puede consistir, por ejemplo, en una serie de mediciones que nos permitan obtener información sobre el sistema. Tales mediciones pueden afectar en mayor o menor grado al sistema (no es lo mismo determinar la masa de un trozo de grafito mediante una balanza que quemarlo y medir el volumen de CO_2 desprendido) pero, así como en mecánica clásica se supone que podemos minimizar tanto como queramos dicha perturbación, esto no siempre es posible en mecánica cuántica. No obstante, las mediciones a las que nos referiremos en lo que sigue serán *ideales* en el sentido de que perturben lo mínimo posible el sistema y *filtrantes* para un valor o un pequeño conjunto de valores (como ocurre siempre que medimos una magnitud continua) de la propiedad u "observable" medido.

Decimos que dos observables de un sistema son *compatibles* cuando, en cualquier estado del sistema en el que esté determinado uno de ellos, una medición ideal del otro no afecta al valor previamente determinado del primero, lo cual puede comprobarse repitiendo la medición del segundo observable*. Un ejemplo de observables no compatibles para una partícula con spin

* Esta afirmación solo es estrictamente cierta si nos limitamos a considerar espacios de Hilbert coherentes, en cada uno de los cuales toman valores únicos ciertos observables, como la carga eléctrica, que se conocen como observables de superselección.

no nulo son las componentes s_z y s_x del spin, ya que la medición de s_x sobre un sistema con $s_z = \hbar/2$ produce un estado en el que pueden obtenerse los valores $\pm\hbar/2$ con igual probabilidad al medir de nuevo s_z (vease ejercicio 1.17). Es evidente que una forma adecuada de preparar un estado será la medición de un conjunto de observables compatibles entre sí. Decimos que dicho conjunto es *maximal* cuando cualquier otro observable compatible con todos los del conjunto es función de ellos. La medición de un conjunto maximal de observables proporciona una información máxima sobre el sistema, en el sentido de que cualquier otra propiedad compatible con las de aquel conjunto estará determinada sin necesidad de realizar nuevas mediciones. Si, además, las mediciones realizadas son filtrantes, el sistema quedará "preparado" en un estado del cual disponemos del máximo grado de información posible, y diremos que hemos preparado el sistema en un estado *puro*. De acuerdo con el presente postulado, el estado del sistema así preparado se representa mediante un vector del espacio de Hilbert asociado. Asimismo, cada rayo unitario de dicho espacio representará, un estado puro del sistema[†].

Si el conjunto de observables medidos no es maximal existirán otros observables compatibles con aquéllos que no quedarán determinados, de modo que podrán existir distintos vectores que correspondan a los valores obtenidos en las mediciones pero difieran en los observables no determinados. En cada caso, será la práctica la que nos permita decidir si un conjunto de observables compatibles es maximal para un sistema dado, ya que esto dependerá del tipo de propiedades y fenómenos que vayamos a considerar e incluso del grado de precisión con que vayamos a trabajar. Por ejemplo, para un átomo de hidrógeno, los observables energía total (H), l^2 y l_z constituyen un conjunto maximal cuando estudiamos fenómenos en los que no se ponen de manifiesto los spines del electrón y del protón, pero será preciso incluir estos observable si queremos analizar experimentos en los que cobren importancia los efectos magnéticos. Por otra parte, muchos aspectos de la resonancia magnética nuclear pueden comprenderse considerando únicamente los observables de spin nucleares; es decir, sin incluir explícitamente en el tratamiento ni la estructura interna del núcleo ni la estructura electrónica. En cambio, ésta última adquiere un papel preponderante en la espectroscopía electrónica.

Un conjunto de observables compatibles maximal *no redundante* (ningún observable del conjunto es función de los restantes) recibe el nombre de *conjunto completo de observables compatibles (CCOC)* y, en general, no es único para un sistema dado.

La medición de los observables de un CCOC es suficiente para preparar un estado puro e, inversamente, en todo estado puro estarán bien definidos los observables de algún CCOC. No obstante, en general, no será fácil idear un dispositivo experimental que permita medir los observables necesarios para preparar el estado puro que corresponde a un vector arbitrario del espacio de Hilbert. Normalmente se elige el CCOC a partir de los observables que más significación tengan en nuestro problema. Por ejemplo, la energía total juega un papel relevante

[†] No consideraremos observables de superselección, que limitan la validez de la afirmación anterior.

en muchos problemas y, completada con otras propiedades que permitan distinguir entre distintos estados de un mismo nivel energético, dará lugar a un CCOC adecuado para tratarlos. Por otra parte, un CCOC que, como veremos en el capítulo 2, tiene especial importancia en relación con la representación de los vectores de estado en una base es el formado por las coordenadas de posición (y, si se requiere, las de spin) de todas las partículas del sistema (en la sección 2.3 se justificará el que tales observables formen un CCOC).

Si no conocemos los valores que toman todos los observables de un CCOC y, de los observables no determinados, sólo disponemos de probabilidades de encontrar cada valor posible, diremos que el sistema se encuentra en un estado *mezcla* (o en una mezcla estadística de estados puros). De acuerdo con esto, no podremos asociar un vector de estado a un estado mezcla, pero podremos caracterizarlo mediante una distribución de probabilidades de un conjunto de vectores de estado.

Ejercicio 1.1. a) Indica como podrías preparar un átomo de hidrógeno en el estado puro correspondiente a uno de los vectores $\psi_{\pm} = (1s\alpha \pm 1s\beta) / \sqrt{2}$ (desprecia los efectos relativistas y comprueba que s_x está bien definido en los estados representados por ψ_{\pm} , en el sentido que se indica en el enunciado del ejercicio 1.7).

b) ¿Cómo podrías preparar un estado con energía $-1/2$ hartree y 50% de probabilidades de obtener $\pm 1/2$ u. a. al medir s_z ? ¿Estará bien definido s_x en dicho estado?

Ejercicio 1.2. Indica un CCOC para el estado interno de un átomo polielectrónico a) sin considerar interacción spin-órbita (en cuyo caso la energía total es compatible con L^2 , L_z , S^2 y S_z), b) considerandola (ahora la energía es compatible con L^2 , S^2 , J^2 y J_z , pero no con L_z ni con S_z) y c) cuando sometemos el átomo a un campo magnético uniforme (que rompe completamente la degeneración energética).

1.2 Segundo postulado: Correspondencia observable \leftrightarrow operador

A cada observable de un sistema físico se le hace corresponder un operador lineal autoadjunto que actúa sobre elementos del espacio de Hilbert asociado al sistema. A las coordenadas cartesianas de posición de las partículas que constituyen el sistema q_1, q_2, \dots, q_N y a sus momentos conjugados p_1, p_2, \dots, p_N , se asocian operadores autoadjuntos independientes del tiempo que satisfagan las reglas de conmutación:

$$[\hat{q}_i, \hat{q}_j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0 \quad \text{y} \quad [\hat{q}_i, \hat{p}_j] = i\hbar \delta_{ij} \hat{1}, \quad \forall i, j = 1, 2, \dots, N.$$

A todo observable cuya expresión clásica se pueda poner en la forma $A(\dots q_i, \dots p_j, \dots t)$ se asocia el operador obtenido substituyendo en esta expresión las coordenadas y momentos por sus operadores cuánticos y simetrizando, si fuera preciso, el operador resultante. Si existen

observables que no tienen análogo clásico, los operadores correspondientes deberán construirse de manera que reproduzcan los hechos experimentales que han conducido a postularlos.

Comentarios:

i) El conmutador de dos operadores \hat{A} y \hat{B} es un nuevo operador definido como $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$. Cuando este operador es nulo ($\hat{A}\hat{B} = \hat{B}\hat{A}$) decimos que \hat{A} y \hat{B} conmutan.

ii) Con la excepción del operador de inversión temporal, los operadores que aparecen en mecánica cuántica son todos lineales por lo que omitiremos este calificativo en lo que sigue.

iii) Dado un operador \hat{A} , su operador *adjunto* \hat{A}^\dagger se define mediante la relación $\langle \psi | \hat{A} \psi' \rangle = \langle \hat{A}^\dagger \psi | \psi' \rangle \quad \forall \psi, \psi' \in \mathcal{H}$; en espacios de dimensión infinita puede ocurrir que \hat{A} y \hat{A}^\dagger tengan dominios distintos (se demuestra que $D(\hat{A}^\dagger) \supseteq D(\hat{A})$). Decimos que un operador es autoadjunto cuando coincide con su adjunto ($\hat{A} = \hat{A}^\dagger$); si esta coincidencia sólo se produce en $D(\hat{A})$ se dice que el operador es hermítico.

iv) La simetrización de la expresión clásica de un observable se precisa, por ejemplo, cuando aquélla contiene el producto de una coordenada y su momento conjugado; en efecto, dado el observable $A = q_i p_i$, no podemos asociarle el operador $\hat{A} = \hat{q}_i \hat{p}_i$ ni el $\hat{A} = \hat{p}_i \hat{q}_i$, ya que ninguno de ellos es autoadjunto, en cambio sí lo es el operador $\hat{A} = (\hat{q}_i \hat{p}_i + \hat{p}_i \hat{q}_i) / 2$.

Ejercicio 1.3. Comprueba las afirmaciones hechas en el comentario iv).

v) Entre los observables sin análogo clásico cabe destacar el spin. Experimentos como el de Stern-Gerlach ponen de manifiesto que en un átomo de hidrógeno en su nivel fundamental, por ejemplo, los observables H , l^2 y l_z no determinan completamente el estado del sistema: hay un observable compatible con ellos que no es función de ellos, ya que puede tomar dos valores una vez fijados los valores que toman los tres primeros. Los experimentos revelan que el nuevo observable tiene mucho en común con el momento angular orbital (carácter vectorial, relación entre sus componentes cartesianas, relación con el momento magnético, etc.) pero es esencialmente distinto de aquel (carácter intrínseco de su módulo, valores que puede tomar, etc.). Los operadores de spin se introducen de manera que reproduzcan tales propiedades (aunque la mecánica cuántica relativista conduce a ellos de manera natural):

$$[\widehat{s}_x, \widehat{s}_y] = i\hbar \widehat{s}_z \quad (+ 2 \text{ relaciones obtenidas permutando cíclicamente } x, y \text{ y } z).$$

1.3 Tercer postulado: Resultados de medidas

Si en un sistema físico que se encuentra en el instante t en un estado puro descrito por el vector normalizado Ψ_t medimos el observable A en aquel instante,

i) la probabilidad de obtener el valor a_i del espectro discreto de \widehat{A} como resultado de la medida es:

$$P_t(A=a_i) = \langle \Psi_t | \widehat{P}_{a_i} | \Psi_t \rangle,$$

ii) la densidad de probabilidad de obtener el valor a del espectro continuo de \widehat{A} como resultado de la medida es:

$$\frac{dP_t(A=a)}{da} = \langle \Psi_t | \widehat{P}_a | \Psi_t \rangle,$$

iii) la probabilidad de obtener un valor que no pertenezca al espectro de \widehat{A} como resultado de la medida es nula.

Comentarios:

i) Daremos una definición de espectro (σ) de un operador que, si bien no es rigurosa matemáticamente, nos servirá para hacernos una idea simple de su significado. Llamaremos espectro de un operador al conjunto formado por sus valores propios en un sentido amplio que enseguida precisaremos. Si el operador es autoadjunto, su espectro es la unión de dos conjuntos disjuntos de números reales: el espectro discreto (σ_d) y el espectro continuo (σ_c). El primero está formado por sus valores propios en sentido estricto, es decir, por los números a_i que cumplen la ecuación

$$\widehat{A} \psi_i = a_i \psi_i$$

con $\psi_i \in \mathcal{H}$, los cuales forman un conjunto numerable. El espectro continuo es, en cambio, un conjunto continuo de valores 'a' que cumplen formalmente una ecuación similar

$$\widehat{A} \psi_a = a \psi_a$$

pero en la que ψ_a no es un vector normalizable del espacio de Hilbert \mathcal{H}^* . No obstante, es posible tratar el espectro continuo de forma análoga a como se trata el discreto utilizando la *delta de Dirac*, función* de la variable 'a' que juega un papel análogo al de la delta de Kro-

* Puede ampliarse la definición de espacio de Hilbert para dar cabida a tales vectores, pero no consideraremos esta posibilidad.

* Este tipo de "funciones" con comportamientos atípicos se conocen en matemáticas con el nombre de *distribuciones*.

necker en el caso discreto. En efecto, la extensión del concepto de ortonormalidad a "vectores" del espectro continuo se expresará así:

$$\langle \psi_a | \psi_{a'} \rangle = \delta(a-a') ;$$

donde la delta de Dirac $\delta(a-a')$ es una función nula para todo 'a' salvo en $a=a'$ y tal que, para cualquier función $f(a)$ continua en 'a',

$$\int_{-\infty}^{\infty} \delta(a-a') f(a) da = f(a') .$$

Es fácil demostrar que el espectro de un operador autoadjunto es real, y que dos vectores propios de un operador autoadjunto con valores propios diferentes son ortogonales entre sí. Si éstos corresponden al espectro continuo la ortogonalidad se expresa mediante la delta de Dirac.

Admitiremos, sin justificarlo, que *los vectores propios de los operadores autoadjuntos* (incluyendo, en su caso, los que corresponden al espectro continuo) *forman un conjunto completo* en el sentido indicado en el comentario i) del primer postulado, es decir, que cualquier elemento del espacio de Hilbert puede expresarse en la forma:

$$\psi = \sum_{a_i \in \sigma_d} \sum_{j=1}^{d_i} c_{ij} \psi_{ij} + \int_{a \in \sigma_c} \sum_{j=1}^{d_a} c_j(a) \psi_{aj} da ,$$

siendo ψ_{ij} un vector propio de \hat{A} con valor propio a_i del espectro discreto de \hat{A} y ψ_{aj} un vector correspondiente al elemento 'a' del espectro continuo de \hat{A} , y siendo d_i y d_a las degeneraciones de a_i y de 'a', respectivamente[†]. La demostración del resultado anterior es trivial para espacios de dimensión finita, pero su extensión a espacios de dimensión infinita excede el nivel de este texto.

Ejercicio 1.4. Indica si los siguientes operadores tienen espectro discreto, continuo o parte discreto y parte continuo, y la degeneración de los valores del espectro: a) hamiltoniano de una partícula libre monodimensional sin spin, b) hamiltoniano del movimiento interno de un átomo de hidrógeno, c) operador momento de una partícula sin spin y d) operador posición de un electrón.

ii) Dado un subespacio \mathcal{H}_1 de un espacio de Hilbert \mathcal{H} , llamamos complemento ortogonal \mathcal{H}_1^\perp de \mathcal{H}_1 en \mathcal{H} al conjunto formado por los vectores de \mathcal{H} ortogonales a todos los vectores de \mathcal{H}_1 . Se demuestra que \mathcal{H}_1^\perp es un subespacio de \mathcal{H} y que todo vector de \mathcal{H} se puede descomponer de forma única como suma de un elemento de \mathcal{H}_1 y uno de su complemento ortogonal:

$$\psi = \psi_1 + \psi_2 \quad \text{con} \quad \psi_1 \in \mathcal{H}_1 \quad \text{y} \quad \psi_2 \in \mathcal{H}_1^\perp .$$

La aplicación \hat{P} que a cada elemento ψ de \mathcal{H} le hace corresponder su componente ψ_1 de \mathcal{H}_1 se

[†] Con el fin de evitar una complicación excesiva de la notación supondremos que los índices de degeneración son siempre discretos.

llama operador de proyección ortogonal o *proyector* de \mathcal{H} sobre \mathcal{H}_1 . Si $\{\dots \phi_i \dots\}$ es una base ortonormal de \mathcal{H}_1 , podemos poner

$$\widehat{P}\psi = \sum_i \phi_i \langle \phi_i | \psi \rangle$$

o, utilizando una notación de evidente carácter mnemotécnico*,

$$\widehat{P} = \sum_i |\phi_i\rangle \langle \phi_i|.$$

En efecto, el vector $\widehat{P}\psi$ pertenece a \mathcal{H}_1 y, completando el conjunto $\{\dots \phi_i \dots\}$ con una base ortonormal de \mathcal{H}_1^\perp , $\{\dots \phi_j' \dots\}$, obtenemos la descomposición única de ψ como suma de un elemento de \mathcal{H}_1 y otro de \mathcal{H}_1^\perp :

$$\begin{aligned} \psi &= \widehat{P}\psi + \sum_j \phi_j' \langle \phi_j' | \psi \rangle \\ &= \widehat{P}\psi + (\widehat{1} - \widehat{P})\psi. \end{aligned}$$

Ejercicio 1.5. Demuestra las siguientes afirmaciones:

- Un operador idempotente ($\widehat{A}^2 = \widehat{A}$) tiene, como máximo, dos valores propios que son 1 y 0.
- Un operador es proyector si y solo si es autoadjunto e idempotente.
- $\langle \psi | \widehat{P}\psi \rangle = \|\widehat{P}\psi\|^2$.
- Una combinación lineal de operadores autoadjuntos es un operador autoadjunto si y sólo si son reales todos los coeficientes de la combinación.
- Una combinación de proyectores sobre subespacios ortogonales es un proyector si y sólo si todos sus coeficientes valen la unidad.

La notación \widehat{P}_{a_i} designará el operador de proyección sobre el subespacio propio de \widehat{A} con valor propio a_i :

$$\widehat{P}_{a_i} \Psi_t = \sum_{j=1}^{d_i} \psi_{ij} \langle \psi_{ij} | \Psi_t \rangle,$$

siendo $\widehat{A} \psi_{ij} = a_i \psi_{ij}$ con $j = 1, 2, \dots, d_i$.

De forma análoga,

$$\widehat{P}_a \Psi_t = \sum_{j=1}^{d_a} \psi_{aj} \langle \psi_{aj} | \Psi_t \rangle,$$

siendo $\widehat{A} \psi_{aj} = a \psi_{aj}$ con $j = 1, 2, \dots, d_a$.

Ejercicio 1.6. Dado el desarrollo de un vector de estado ψ en una base de estados propios de un

* La notación "bra-ket", introducida por Dirac y muy extendida en mecánica cuántica, consiste en representar mediante $|\psi\rangle$ los vectores del espacio de Hilbert ("kets") y $\langle \psi|$ los elementos correspondientes del espacio dual ("bras"). Dado que las nociones de vector de estado y producto escalar son suficientes para desarrollar la presente exposición, no utilizaremos el concepto de "bra" y consideraremos la expresión general $|\psi_i\rangle \langle \psi_j|$ como notación mnemotécnica para designar el operador que transforma el vector ψ en el vector $\psi_i \langle \psi_j | \psi \rangle$. En adelante utilizaremos indistintamente las notaciones ψ y $|\psi\rangle$ para los vectores de estado.

observable A:

$$\Psi = \sum_{a_i \in \sigma_d} \sum_{j=1}^{d_i} c_{ij} \psi_{ij} + \int_{a \in \sigma_c} \sum_{j=1}^{d_a} c_j(a) \psi_{aj} da ,$$

demuestra que

$$P_{\Psi}(A=a_i) = \sum_{j=1}^{d_i} |c_{ij}|^2$$

y

$$dP_{\Psi}(A=a) / da = \sum_{j=1}^{d_a} |c_j(a)|^2 .$$

Los coeficientes c_{ij} y $c_j(a)$ cuyos módulos elevados al cuadrado representan probabilidades y densidades de probabilidad, se denominan, respectivamente, *amplitudes de probabilidad* y *amplitudes de densidad de probabilidad*.

Ejercicio 1.7. Demuestra que un observable A está bien definido en un estado puro Ψ (es decir, $P_{\Psi}(A=a) = 1$) si y sólo si Ψ es propio de \hat{A} .

Ejercicio 1.8. ¿Qué valores pueden obtenerse y con qué probabilidades al medir, en el instante $t = 0$, la energía total y la componente z de los momentos angulares orbital y de spin del electrón de un átomo de hidrógeno descrito por el vector de estado $\Psi_{t=0} = \frac{\sqrt{3}}{2} \phi_{1s} \alpha + \frac{1}{2} e^{i\pi/4} \phi_{2p_0} \beta$?

iii) Todo operador autoadjunto se puede expresar mediante su *descomposición espectral*:

$$\hat{A} = \sum_{a_i \in \sigma_d} a_i \hat{P}_{a_i} + \int_{a \in \sigma_c} a \hat{P}_a da .$$

Ejercicio 1.9. Deduce la relación anterior teniendo en cuenta el comentario i).

Ejercicio 1.10. Demuestra que, para cualquier sistema, existe un operador (o más de uno) que, por si solo, constituye un CCOC. *Sugerencia:* construye el operador buscado a partir de su descomposición espectral. Puedes encontrar un ejemplo en el apartado c) del ejercicio 1.2.

Dada una función de variable real $f(x)$, definimos el operador "función f de \hat{A} " ($f(\hat{A})$) como aquel cuya descomposición espectral es

$$f(\hat{A}) = \sum_{a_i \in \sigma_d} f(a_i) \hat{P}_{a_i} + \int_{a \in \sigma_c} f(a) \hat{P}_a da .$$

Ejercicio 1.11. En un átomo polielectrónico y prescindiendo de efectos relativistas, ¿es posible expresar los operadores \hat{L}^2 y \hat{S}^2 como funciones de \hat{H} ?

El operador identidad se puede considerar como función de cualquier otro, por lo que podemos escribir una descomposición espectral o *resolución de la identidad* a partir del espectro de cualquier operador:

$$\hat{1} = \sum_{a \in \sigma_d} \hat{P}_{a_i} + \int_{a \in \sigma_c} \hat{P}_a da .$$

Si $f(x)$ es una función analítica de x , podemos desarrollar $f(a_i)$ y $f(a)$ en serie de Taylor, lo cual permite expresar $f(\hat{A})$ en la forma

$$f(\hat{A}) = f(0) + f'(0) \hat{A} + f''(0) \hat{A}^2 / 2! + f'''(0) \hat{A}^3 / 3! + \dots .$$

iv) Llamamos *valor esperado* del observable A en el estado Ψ_t ($\langle A \rangle_{\Psi_t}$) al promedio de los resultados que se obtendrían al medir el A en infinitos sistemas idénticos todos ellos en el estado Ψ_t . De acuerdo con el presente postulado,

$$\langle A \rangle_{\Psi_t} = \sum_{a \in \sigma_d} a_i \langle \Psi_t | \hat{P}_{a_i} \Psi_t \rangle + \int_{a \in \sigma_c} a \langle \Psi_t | \hat{P}_a \Psi_t \rangle da$$

o, introduciendo la descomposición espectral de \hat{A} ,

$$\boxed{\langle A \rangle_{\Psi_t} = \langle \Psi_t | \hat{A} \Psi_t \rangle} .$$

v) Dado que, en general, la medición de un observable A en sistemas idénticos y preparados en el mismo estado Ψ no conduce a un único resultado, interesa disponer de una medida del grado de dispersión que presentan los valores obtenidos. Esta medida se define usualmente como la desviación cuadrática media o *dispersión* del observable A en el estado Ψ :

$$\begin{aligned} \Delta_{\Psi} A &\equiv \langle \Psi | (\hat{A} - \langle A \rangle_{\Psi})^2 \Psi \rangle^{1/2} \\ &= \{ \langle A^2 \rangle_{\Psi} - (\langle A \rangle_{\Psi})^2 \}^{1/2} . \end{aligned}$$

Dados dos observables A y B de un mismo sistema, $\Delta_{\Psi} A$ y $\Delta_{\Psi} B$ cumplen la siguiente *relación de dispersión*:

$$\Delta_{\Psi} A \Delta_{\Psi} B \geq | \langle \Psi | [\hat{A}, \hat{B}] \Psi \rangle | / 2 .$$

Para demostrarlo definamos los operadores autoadjuntos $\hat{A}' \equiv \hat{A} - \langle A \rangle_{\Psi}$ y $\hat{B}' \equiv \hat{B} - \langle B \rangle_{\Psi}$.

Si λ es un parámetro real,

$$\langle \Psi | (\hat{A}' - i\lambda \hat{B}') (\hat{A}' + i\lambda \hat{B}') \Psi \rangle = \| (\hat{A}' + i\lambda \hat{B}') \Psi \|^2 \geq 0 ,$$

es decir,

$$\langle \Psi | \hat{B}'^2 \Psi \rangle \lambda^2 + \langle \Psi | i[\hat{A}', \hat{B}'] \Psi \rangle \lambda + \langle \Psi | \hat{A}'^2 \Psi \rangle \geq 0$$

y, teniendo en cuenta que $[\hat{A}', \hat{B}'] = [\hat{A}, \hat{B}]$ y la definición de dispersión de un observable,

$$(\Delta_{\Psi} B)^2 \lambda^2 + \langle \Psi | i[\hat{A}, \hat{B}] \Psi \rangle \lambda + (\Delta_{\Psi} A)^2 \geq 0 .$$

El primer miembro de esta desigualdad es un polinomio de segundo grado en λ con coeficientes

reales, ya que

$$\langle \Psi | i[\hat{A}, \hat{B}] \Psi \rangle = \langle -i[\hat{B}, \hat{A}] \Psi | \Psi \rangle = \langle i[\hat{A}, \hat{B}] \Psi | \Psi \rangle$$

por lo que podemos utilizar el resultado del ejercicio 1.12 para establecer la desigualdad

$$4 (\Delta_{\Psi} A)^2 (\Delta_{\Psi} B)^2 - \langle \Psi | i[\hat{A}, \hat{B}] \Psi \rangle^2 \geq 0,$$

que equivale a la relación de dispersión.

Ejercicio 1.12. Demuestra que el polinomio de segundo grado en la variable real x con coeficientes reales $ax^2 + bx + c$ y $a > 0$ toma valores ≥ 0 para todo x si y sólo si $4ac - b^2 \geq 0$.

Ejercicio 1.13. Demuestra que, si el conmutador de los operadores \hat{A} y \hat{B} asociados a dos observables A y B es una constante no nula, no existen estados con ambos observables bien definidos. ¿Pueden existir tales estados cuando el conmutador no es una constante?

vi) Si un sistema se halla, en el instante t , en un estado mezcla con probabilidades $p_1, p_2, \dots, p_i, \dots$ para los vectores normalizados independientes $\psi_1, \psi_2, \dots, \psi_i, \dots$ (con $0 \leq p_1, p_2, \dots \leq 1$ y $\sum_i p_i = 1$), el valor esperado de un observable A será un promedio ponderado de los valores esperados que corresponderían a cada uno de dichos vectores:

$$\langle A \rangle_t = \sum_i p_i \langle A \rangle_{\psi_i}.$$

Introduciendo el *operador densidad*

$$\hat{\rho}_t = \sum_i p_i |\psi_i\rangle \langle \psi_i|$$

para representar matemáticamente aquel estado mezcla, y un conjunto numerable denso ortonormal $\{\dots \phi_i \dots\}$ podemos poner:

$$\begin{aligned} \langle A \rangle_{\rho_t} &= \sum_i p_i \langle \psi_i | \hat{A} \psi_i \rangle \\ &= \sum_i \sum_j p_i \langle \psi_i | \phi_j \rangle \langle \phi_j | \hat{A} \psi_i \rangle \\ &= \sum_i \sum_j p_i \langle \phi_j | \hat{A} | \psi_i \rangle \langle \psi_i | \phi_j \rangle \\ &= \sum_j \langle \phi_j | \hat{A} \hat{\rho}_t \phi_j \rangle, \end{aligned}$$

y, definiendo la *traza* de un operador \hat{B} como

$$\text{Tr } \hat{B} \equiv \sum_j \langle \phi_j | \hat{B} \phi_j \rangle$$

$$\langle A \rangle_{\rho_t} = \text{Tr} (\hat{A} \hat{\rho}_t).$$

Ejercicio 1.14. Comprueba las siguientes afirmaciones.

a) La traza de un operador es independiente del conjunto numerable denso ortonormal utilizado

* En la sección 2.1 veremos que la traza de un operador es la traza de la matriz que lo representa en una base ortonormal.

para calcularla (utiliza la propiedad $\sum_i L_{ki} L_{ji}^* = \delta_{kj}$ que cumple la matriz de cambio de una base ortonormal a otra, y que deducirás en el ejercicio 2.2).

b) La traza de un operador hermítico es la suma de sus valores propios.

c) $\text{Tr}(\widehat{A}\widehat{B}) = \text{Tr}(\widehat{B}\widehat{A})$.

d) $\text{Tr}(\widehat{A}_1 \dots \widehat{A}_n) = \text{Tr}(\widehat{A}_{i_1} \dots \widehat{A}_{i_n})$.

donde $i_1 \dots i_n$ es cualquier permutación cíclica de los índices $1 \dots n$.

e) $\text{Tr}\widehat{\rho} = 1$.

f) $\widehat{\rho}^\dagger = \widehat{\rho}$.

Ejercicio 1.15. Demuestra que, para cualquier operador \widehat{A} y cualquier vector Ψ ,

$$\text{Tr}(\widehat{A}|\Psi\rangle\langle\Psi|) = \langle\Psi|\widehat{A}|\Psi\rangle .$$

Ejercicio 1.16. Demuestra que la probabilidad (o la densidad de probabilidad) de obtener un valor 'a' al medir el observable A en un sistema que se encuentra en el estado descrito por el operador densidad $\widehat{\rho}$ en el momento de realizar la medición viene dada por

$$P_\rho(A=a) = \text{Tr}(\widehat{P}_a \widehat{\rho}) = \sum_{j=1}^{d_a} \langle \psi_{aj} | \widehat{\rho} | \psi_{aj} \rangle .$$

Un estado es puro si y sólo si su operador densidad es un proyector sobre un subespacio unidimensional. En efecto, un estado puro puede representarse mediante una mezcla estadística de estados con probabilidad 1 para uno de ellos y 0 para los restantes, de modo que podemos asociarle un operador densidad igual al proyector sobre el vector del estado que tiene probabilidad 1. Recíprocamente, si el operador densidad es un proyector cumplirá la relación de idempotencia $\widehat{\rho}^2 = \widehat{\rho}$, y sus valores propios serán 1 o 0 (ejercicio 1.5 a)). Como la traza de un operador es la suma de sus valores propios y $\text{Tr}\widehat{\rho} = 1$ (ejercicio 1.14 b) y e)), $\widehat{\rho}$ tendrá un vector propio con valor propio 1 y los demás tendrán valor propio 0; es decir, habrá un estado puro con probabilidad 1.

Los operadores densidad constituyen, pues, una alternativa a los vectores de estado para representar estados puros que permite tratar a éstos como caso particular de los estados mezcla. Este formalismo no presenta la ambigüedad que surge al representar los estados puros mediante vectores de estado, ya que los distintos vectores de estado correspondientes a un mismo rayo unitario conducen a un mismo operador de proyección.

1.4 Cuarto postulado: Colapso del estado

Dado un sistema físico que, en el instante t , se encuentra en el estado descrito por el operador densidad $\hat{\rho}_t$, si realizamos en dicho instante una medida ideal del observable A filtrante para un conjunto de valores $\Delta = \Delta_d \cup \Delta_c$, Δ_d del espectro discreto de \hat{A} y Δ_c de su espectro continuo, el sistema queda en el estado descrito por el operador densidad

$$\hat{\rho}_\Delta = \frac{\hat{P}_\Delta \hat{\rho}_t \hat{P}_\Delta}{\text{Tr}(\hat{P}_\Delta \hat{\rho}_t)} .$$

siendo $\hat{P}_\Delta \equiv \sum_{a \in \Delta_d} \hat{P}_a + \int_{a \in \Delta_c} \hat{P}_a da$.

Comentarios:

i) En general, la medición (aunque sea ideal) cambia el estado del sistema. No obstante, tal alteración no se produce cuando \hat{P}_Δ conmuta con $\hat{\rho}_t$. Esta caso será discutido en el comentario ix).

ii) En general, el estado del sistema se modifica al efectuar la medición. El estado resultante, si bien no será en general puro, tiene probabilidad 1 para los valores del observable A contenidos en Δ . En efecto, $\hat{P}_\Delta \hat{\rho}_t \hat{P}_\Delta$ es la proyección de $\hat{\rho}_t$ sobre el subespacio suma directa* de los subespacios propios de \hat{A} correspondientes a valores de Δ , y el denominador que aparece en la definición de $\hat{\rho}_\Delta$ tiene por objeto normalizar aquella proyección a traza unidad:

$$\begin{aligned} \text{Tr}(\hat{P}_\Delta \hat{\rho}_t \hat{P}_\Delta) &= \text{Tr}(\hat{P}_\Delta \hat{P}_\Delta \hat{\rho}_t) \\ &= \text{Tr}(\hat{P}_\Delta \hat{\rho}_t) , \end{aligned}$$

de modo que, para un conjunto Δ discreto (vease ejercicio 1.16),

$$\begin{aligned} P_{\rho_\Delta}(A \in \Delta) &= \sum_{a \in \Delta} P_{\rho_\Delta}(A=a) \\ &= \sum_{a \in \Delta} \text{Tr}(\hat{P}_a \hat{\rho}_\Delta) \\ &= \text{Tr}(\hat{P}_\Delta \hat{\rho}_\Delta) \\ &= 1 , \end{aligned}$$

resultado generalizable a conjuntos Δ con parte continua.

Si la medida no fuera ideal el estado resultante dependería, en general del aparato de medida utilizado y, por lo tanto, no podría determinarse conociendo únicamente $\hat{\rho}_t$ y el conjunto de

* Decimos que un espacio vectorial es suma directa de dos o más subespacios cuando cada elemento del primero se puede expresar de forma unívoca como suma de un elemento de cada subespacio, y se designa con el símbolo \oplus . En el comentario ii) de la sección 1.3 vimos un ejemplo: $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1^\perp$. El espacio de los números complejos es suma directa del de los reales y del de los imaginarios puros.

valores filtrados.

iii) Si la medida filtra un *único valor* de A , esta propiedad queda perfectamente definida en el estado resultante y , si dicho valor es *no degenerado*, el sistema queda en el estado *puro* cuyo vector es propio de \hat{A} con valor propio el obtenido en la medida, estado que designaremos mediante la notación abreviada $|a\rangle \equiv |\psi_a\rangle$:

$$\begin{aligned} \hat{\rho}_a &= \frac{|a\rangle\langle a| \sum_i p_i |\psi_i\rangle\langle\psi_i| |a\rangle\langle a|}{\sum_j \langle\phi_j||a\rangle\langle a|\sum_i p_i |\psi_i\rangle\langle\psi_i|\phi_j\rangle} \\ &= \frac{|a\rangle\langle a| \sum_i p_i |\langle\psi_i|a\rangle|^2}{\sum_i p_i |\langle\psi_i|a\rangle|^2} \\ &= |a\rangle\langle a|, \end{aligned}$$

resultado que es válido tanto si la medida filtrante es ideal como si no lo es. Observemos que el estado resultante de la medida es, en este caso, independiente del estado de partida.

iv) Si el sistema estaba en un estado *puro* Ψ inmediatamente antes de realizar la medición, el estado que resulta es también puro, independientemente de que se filtre uno o más valores y de la degeneración de éstos. En efecto, teniendo en cuenta el resultado del ejercicio 1.15,

$$\begin{aligned} \hat{\rho}_\Delta &= \frac{\hat{P}_\Delta |\Psi\rangle\langle\Psi| \hat{P}_\Delta}{\text{Tr}(\hat{P}_\Delta |\Psi\rangle\langle\Psi|)} \\ &= \frac{|\hat{P}_\Delta \Psi\rangle\langle\hat{P}_\Delta \Psi|}{\langle\Psi|\hat{P}_\Delta|\Psi\rangle} \\ &= \left| \frac{\hat{P}_\Delta \Psi}{\|\hat{P}_\Delta \Psi\|} \right\rangle \left\langle \frac{\hat{P}_\Delta \Psi}{\|\hat{P}_\Delta \Psi\|} \right|, \end{aligned}$$

es decir, el vector Ψ , que representaba del estado del sistema inmediatamente antes de efectuar la medición, cambiará de la siguiente forma:

$$\boxed{\Psi \rightarrow \frac{\hat{P}_\Delta \Psi}{\|\hat{P}_\Delta \Psi\|}}.$$

v) Cuando filtramos valores en una zona Δ continua del espectro, es imposible seleccionar un único valor del observable medido. No obstante, podemos suponer que estamos midiendo el observable discretizado \hat{P}_Δ , para el cual el valor obtenido es único (si bien infinitamente degenerado). En cualquier caso, la medición de observables continuos no permitirá, en general, preparar estados puros a partir de estados mezcla. Un ejemplo típico de este caso se presenta cuando hacemos pasar una partícula por una rendija de anchura Δ para obtener información sobre la coordenada de posición (x) perpendicular a la dirección de la rendija. Si la partícula atraviesa la rendija quedará en un estado con valor 1 para el observable discretizado "haber

pasado (valor 1) o no (valor 0) por la rendija":

$$\widehat{P}_\Delta \equiv \int_{x \in \Delta} \widehat{P}_x dx .$$

En realidad, si el conjunto Δ está inmerso en un conjunto continuo más amplio, sus límites no estarán bien definidos y el integrando de la expresión anterior debería incluir una función moduladora $f(x)$ que decaiga más o menos rápidamente a cero en la frontera de Δ . La forma de esta función dependerá del dispositivo experimental utilizado y, de acuerdo con el carácter ideal de las medidas que consideraremos, no será considerada en lo que sigue.

Observables del tipo \widehat{P}_Δ pueden introducirse también para conjuntos Δ que contengan partes discretas.

vi) Un caso límite trivial se presenta cuando el conjunto Δ coincide con el espectro del operador asociado al observable medido. Podemos, entonces, suponer que hemos medido el observable identidad y, por lo tanto, no se produce colapso del estado. Este sería el caso de un aparato de Stern-Gerlach en el que se separa en dos un haz de partículas de spin 1/2 (también puede considerarse una única partícula) y, sin que quede registrada ninguna señal detectable de la ruta seguida por cada partícula, se reconducen los dos haces obtenidos para obtener, de nuevo, un único haz.

vii) Supondremos que el postulado es aplicable a mediciones ideales no filtrantes siempre que quede un registro macroscópico detectable del valor o los valores resultantes de la medida (independientemente de que algún ser consciente observe dicho registro).

Ejercicio 1.17. El estado de un átomo de hidrógeno de una muestra gaseosa en equilibrio térmico a una temperatura T se puede describir mediante el operador densidad $\widehat{\rho} = (1/Z) \exp(-\widehat{H}/kT)$ (vease ejercicio 1.20). En el supuesto de que los observables H , l^2 y l_z constituyan un CCOC adecuado para describir el estado interno del átomo, indica el estado en el que quedará un átomo de la muestra después de medir su energía interna y obtener

a) el valor $-1/2$ hartree,

b) el valor $-1/8$ hartree.

Si incluimos s_z en nuestro CCOC

c) ¿qué estado se obtendría después de medir la energía con resultado $-1/2$ hartree?

d) ¿qué estado se obtendría si, después de la medición realizada en el punto anterior, sometemos el átomo a una medida de s_x filtrante para $s_x = 1/2$ u. a. (recuerda el ejercicio 1.1)?

Un haz de átomos de hidrógeno preparados tal como se indica en c) y d) atraviesa un dispositivo de Stern-Gerlach con el campo magnético orientado según la dirección del eje z ; a la salida de éste, los dos haces emergentes son reconducidos (según trayectorias simétricas) a la dirección inicial e introducidos en un nuevo dispositivo de Stern-Gerlach orientado según la dirección del

eje x; finalmente, una placa detectora registra el impacto de los átomos que emergen del conjunto. Indica el número de puntos de impacto y la disposición de éstos sobre la placa en los siguientes casos:

- e) si filtramos los átomos con $s_z = 1/2$ u. a. en el primer dispositivo;
- f) si filtramos los átomos con $s_z = -1/2$ u. a. en el primer dispositivo;
- g) si no filtramos los átomos en el primer dispositivo pero registramos su trayectoria a través del mismo;
- h) si no filtramos los átomos en el primer dispositivo ni registramos su trayectoria a través de éste y del sistema reductor;
- i) si no filtramos los átomos en el primer dispositivo ni registramos su trayectoria, y existe una diferencia de longitud entre las trayectorias correspondientes a $s_z = 1/2$ u. a. y $s_z = -1/2$ u. a. que implica un diferencia de 2π segundos en el tiempo invertido por un átomo en recorrerlas.

Ejercicio 1.18. Demuestra que un estado mezcla con probabilidades de 1/2 para los estados puros $1s\alpha$ y $1s\beta$ es indistinguible de uno con probabilidades 1/2 para estados con $s_x = 1/2$ u. a. y $s_x = -1/2$ u. a..

viii) Si A y B son observables compatibles para un sistema dado, las medidas sucesivas (o simultáneas) de uno y otro producen estados en los que ambos observables están bien determinados; realizando dichas medidas sobre distintos estados puros, podemos construir un conjunto completo de estados puros propios de ambos observables. Se demuestra que esto es condición necesaria y suficiente para que conmuten los correspondientes operadores. Extendiendo este razonamiento a los observables de un CCOC, concluimos que todos ellos han de conmutar entre si y que ha de existir un conjunto completo de estados propios de todos ellos a la vez. Dicho conjunto es único (salvo fases) ya que, si hubiera dos bases diferentes $\{\dots \psi_i \dots\}$ y $\{\dots \phi_i \dots\}$ con algún ψ_i no contenido en la segunda, podríamos expresar dicho vector como combinación lineal de dos o más vectores de la segunda base que diferirían en algún valor de alguno de los observables del CCOC, y ψ_i no sería propio de estos.

ix) Si el operador asociado al observable medido conmuta con $\hat{\rho}_t$ existirá un conjunto completo de estados propios de ambos operadores. Estos estados también serán propios de \hat{P}_Δ , como se comprueba aplicándoles la definición de este operador dada en el enunciado del postulado. Entonces \hat{P}_Δ y $\hat{\rho}_t$ también conmutan y la medición no altera el estado del sistema. De hecho, basta que \hat{P}_Δ y $\hat{\rho}_t$ conmuten para que esto ocurra.

1.5 Quinto postulado: Evolución libre

Un sistema preparado en un estado puro continua en un estado puro en el intervalo de tiempo que transcurre hasta la realización de nuevas mediciones sobre el sistema; un vector representativo del estado del sistema evoluciona en dicho intervalo de tiempo de acuerdo con la ecuación de Schrödinger:

$$\boxed{i\hbar \frac{d\Psi_t}{dt} = \hat{H}(t) \Psi_t},$$

donde $\hat{H}(t)$ es el hamiltoniano del sistema.

Comentarios:

i) El hamiltoniano es el operador asociado a la función hamiltoniana clásica del sistema, la cual, cuando no depende del tiempo, coincide con la energía total; este es el caso de cualquier sistema aislado y, de hecho solo encontraremos hamiltonianos dependientes del tiempo cuando estudiemos un sistema en interacción con otro que no se desea incluir en el tratamiento cuántico (por ejemplo, una molécula sometida a un campo electromagnético oscilante descrito clásicamente). Un estudio de este tipo es aproximado, ya que, como veremos en la discusión del siguiente postulado, el vector de estado que describe dos sistemas en interacción mutua no es, en general, separable en vectores de estado que representen cada parte. No obstante, el estudio de una de las partes incluyendo en su hamiltoniano los efectos de la otra supone una notable simplificación del problema que, bajo ciertas condiciones, conduce a predicciones correctas.

ii) La evolución temporal de los valores esperados, que es lo que se puede corroborar mediante experimentos, se deduce inmediatamente a partir del presente postulado:

$$\frac{d \langle \Psi_t | \hat{A}(t) \Psi_t \rangle}{dt} = \left\langle \frac{d \Psi_t}{dt} | \hat{A}(t) \Psi_t \right\rangle + \left\langle \Psi_t | \left(\frac{d \hat{A}(t)}{dt} \right) \Psi_t \right\rangle + \left\langle \Psi_t | \hat{A}(t) \frac{d \Psi_t}{dt} \right\rangle$$

y, utilizando la ecuación de Schrödinger,

$$i\hbar \frac{d \langle \Psi_t | \hat{A}(t) \Psi_t \rangle}{dt} = \langle \Psi_t | [\hat{A}(t), \hat{H}(t)] \Psi_t \rangle + i\hbar \left\langle \Psi_t | \left(\frac{d \hat{A}(t)}{dt} \right) \Psi_t \right\rangle.$$

Si el valor esperado de un observable es independiente del tiempo para *cualquier* estado puro del sistema, decimos que dicho observable es una *constante de movimiento*. De acuerdo con la ecuación anterior, A es constante de movimiento de un sistema cuyo hamiltoniano es $\hat{H}(t)$ si y sólo si

$$[\hat{A}(t), \hat{H}(t)] + i\hbar (d \hat{A}(t) / dt) = 0,$$

ecuación que, para observables cuyo operador no depende del tiempo, se reduce a

$$[\hat{A}, \hat{H}(t)] = 0.$$

iii) La evolución que expresa la ecuación de Schrödinger puede representarse mediante un *operador de evolución temporal* ($\hat{U}(t, t')$) definido de forma que se cumpla la relación

$$\Psi(t) = \hat{U}(t, t_0) \Psi(t_0)$$

para cualquier estado puro del sistema. Introduciendo esta relación en la ecuación de Schrödinger se obtiene la ecuación diferencial que determina el operador \hat{U} :

$$i\hbar \frac{d\hat{U}(t, t_0)}{dt} = \hat{H}(t) \hat{U}(t, t_0).$$

Se demuestra que el operador de evolución temporal es unitario ($\hat{U}^{-1} = \hat{U}^\dagger$), lo cual garantiza la constancia de la norma de los vectores de estado en el tiempo.

Si \hat{H} no depende del tiempo la ecuación anterior, junto con la condición inicial $\hat{U}(t_0, t_0) = \hat{1}$, conduce a la solución $\hat{U}(t, t_0) = \exp[-i(t-t_0)\hat{H}/\hbar]$.

iv) En un sistema cuyo hamiltoniano no depende del tiempo los vectores propios de dicho operador representan estados que llamaremos *estacionarios*.

Ejercicio 1.19. Indica como evolucionan en el tiempo

- un vector que represente un estado estacionario y
- el valor esperado en un estado estacionario de un observable cuyo operador no depende del tiempo.

Observemos que la constancia de los valores esperados de una constante de movimiento no depende de que el estado considerado sea o no estacionario.

v) La evolución de un estado *mezcla* se obtiene considerando como evolucionan los estados puros que componen la mezcla estadística:

$$\hat{\rho}_t = \sum_i p_i |\psi_i(t)\rangle \langle \psi_i(t)| = \sum_i p_i |\hat{U}(t, t_0) \psi_i(t_0)\rangle \langle \hat{U}(t, t_0) \psi_i(t_0)|,$$

es decir,

$$\hat{\rho}_t = \hat{U}(t, t_0) \hat{\rho}_{t_0} \hat{U}^\dagger(t, t_0).$$

Derivando esta expresión se obtiene la siguiente ecuación diferencial de evolución:

$$i\hbar \frac{d\hat{\rho}_t}{dt} = [\hat{H}(t), \hat{\rho}_t].$$

Ejercicio 1.20. Indica como evolucionan en el tiempo

- el operador densidad asociado a un estado puro estacionario y

b) un operador densidad que represente una mezcla estadística de estados estacionarios.

Un importante ejemplo de este último caso es el operador densidad que describe un sistema macroscópico aislado en equilibrio termodinámico a una temperatura T:

$$\hat{\rho} = \sum_{ij} \frac{\exp(-E_i/kT)}{\sum_{lm} \exp(-E_l/kT)} |\psi_{ij}\rangle\langle\psi_{ij}| = \frac{1}{Z} \exp(-\hat{H}/kT),$$

donde hemos supuesto que el espectro de \hat{H} es discreto y que $\hat{H} \psi_{ij} = E_i \psi_{ij}$, $j=1, \dots, d_i$.

Ejercicio 1.21. Considera un oscilador armónico unidimensional en un estado mezcla con probabilidad 0,5 para los estados estacionarios fundamental (ϕ_0) y primer excitado (ϕ_1).

- ¿Cuál será el valor esperado de la posición en dicho estado? (no es preciso que efectúes ningún cálculo).
- Comprueba que el estado considerado también se puede definir como mezcla estadística al 50% de los estados puros no estacionarios $\psi_{\pm} = (\phi_0 \pm \phi_1)/\sqrt{2}$. ¿Cómo evoluciona en el tiempo el valor esperado de la posición en cada uno de estos estados puros? ¿Cómo se explica que pueda ser "estacionaria" una mezcla estadística de estados puros no estacionarios?

Ejercicio 1.22. Demuestra que la ecuación de evolución del valor esperado de un observable en un estado mezcla tiene la misma forma que la correspondiente a un estado puro:

$$i\hbar \frac{d}{dt} (\text{Tr } \hat{A} \hat{\rho}) = \text{Tr} ([\hat{A}, \hat{H}] \hat{\rho}) + i\hbar \text{Tr} \{ (\frac{d\hat{A}}{dt}) \hat{\rho} \}.$$

vi) La forma en que hemos enunciado los postulados segundo y quinto no es la única manera de representar la evolución libre (es decir, entre medidas consecutivas) de un sistema en el marco de la mecánica cuántica. En efecto, hemos supuesto que los operadores correspondientes a posiciones y momentos son independientes del tiempo y, lo mismo podemos decir de cualquier otro observable a no ser que el hamiltoniano del sistema incluya términos que dependan explícitamente del tiempo; la evolución libre corre, pues, a cargo de los vectores de estado o de las matrices de densidad, tal como establece la ecuación de Schrödinger. Esta forma de tratar la evolución recibe el nombre de *imagen de Schrödinger*. No obstante, lo que se observa experimentalmente no son vectores de estado o matrices de densidad, sino valores esperados en los que interviene, junto con aquellos, un operador relacionado con algún observable. Por lo tanto, podríamos haber asignado la evolución temporal a dichos operadores, lo que conduciría a una forma alternativa de representar la evolución temporal que se conoce como *imagen de Heisenberg*. La ecuación de evolución en esta imagen se obtiene fácilmente teniendo en cuenta que ha de conducir a los mismos valores esperados que la imagen de Schrödinger en cualquier intervalo de tiempo de evolución libre:

$$\langle \Psi_H | \hat{A}_H(t) | \Psi_H \rangle = \langle \Psi_S(t) | \hat{A}_S(t) | \Psi_S(t) \rangle,$$

donde los subíndices 'H' y 'S' se refieren a las imágenes de Heisenberg y Schrödinger,

respectivamente. En efecto, definiendo

$$\begin{aligned}\Psi_H &\equiv \hat{U}^\dagger(t, t_0) \Psi_S(t) \\ &= \Psi_S(t_0)\end{aligned}$$

y utilizando la igualdad de los valores esperados, se obtiene la ley de transformación para los operadores asociados a los observables en una y otra imagen:

$$\hat{A}_H(t) = \hat{U}^\dagger(t, t_0) \hat{A}_S(t) \hat{U}(t, t_0)$$

(es obvio que el operador de evolución temporal adopta la misma forma en ambas imágenes, por lo que no precisa subíndice).

Derivando la expresión anterior respecto del tiempo y utilizando la ecuación de Schrödinger se obtiene la ecuación de evolución libre en la imagen de Heisenberg:

$$i\hbar \frac{d\hat{A}_H(t)}{dt} = [\hat{A}_H(t), \hat{H}_H(t)] + i\hbar \left(\frac{d\hat{A}_S(t)}{dt} \right)_H .$$

Ejercicio 1.23. Comprueba las siguientes afirmaciones:

- la condición de constante de movimiento en la imagen de Heisenberg es $d\hat{A}_H(t)/dt = 0$;
- si \hat{H}_S es independiente del tiempo, $\hat{H}_H = \hat{H}_S$.

Ejercicio 1.24. Para una partícula unidimensional de masa m sometida a un potencial que solo depende de la posición,

- demuestra que $m d\hat{x}_H/dt = \hat{p}_H$;
- deduce la segunda ley de Newton en la imagen de Heisenberg teniendo en cuenta que, si $[\hat{A}, \hat{B}]$ conmuta con \hat{B} , $[\hat{A}, f(\hat{B})] = [\hat{A}, \hat{B}] df(\hat{B})/d\hat{B}$.

Aunque la imagen de Schrödinger es, con mucho, la más utilizada, la de Heisenberg proporciona una visión más próxima a la imagen clásica, como se aprecia en el ejercicio anterior, y resulta especialmente adecuada para tratar determinados problemas, como es el caso del estudio cuántico de la radiación electromagnética.

Entre las posturas extremas que se adoptan en las dos imágenes consideradas, existe una infinidad de imágenes en las que la evolución temporal se reparte de distinta forma entre vectores de estado y operadores; de hecho, cualquier transformación del tipo $\Psi_T(t) = \hat{V}(t) \Psi_S(t)$ con $\hat{V}(t)$ unitario define una imagen de evolución temporal. La única de estas que tiene cierta importancia es la *imagen de Dirac* o de interacción, que se utiliza en problemas cuyo hamiltoniano es de la forma $\hat{H}(t) = \hat{H}_0 + \hat{H}'(t)$ con $\hat{H}'(t) \ll \hat{H}_0$. En dicha imagen se elimina de los vectores de estado la parte de la evolución debida a \hat{H}_0 , dejando únicamente la debida a $\hat{H}'(t)$:

$$\Psi_D(t) = \exp(i(t-t_0)\hat{H}_0/\hbar) \Psi_S(t) .$$

Dicha parte de la evolución se asigna a los operadores:

$$\hat{A}_D(t) = \exp(i(t-t_0)\hat{H}_0/\hbar) \hat{A}_S(t) \exp(-i(t-t_0)\hat{H}_0/\hbar)$$

Ejercicio 1.25. Deduce las ecuaciones de evolución en la imagen de Dirac para los vectores de estado y para los operadores asociados a observables.

1.6 Sexto postulado: Sistemas compuestos

El espacio de Hilbert asociado a un sistema compuesto de subsistemas es el producto directo de los espacios asociados a cada subsistema. Si dos de dichos subsistemas son idénticos, los vectores de estado del sistema compuesto han de ser simétricos o antisimétricos respecto del intercambio de tales subsistemas según sea entero o semiimpar, respectivamente, el momento angular total de éstos.

Comentarios:

i) Este postulado es un complemento del primero que, por razones pedagógicas, suele enunciarse separadamente.

ii) Sean \mathcal{H}_1 y \mathcal{H}_2 dos espacios vectoriales definidos sobre el mismo cuerpo escalar. Llamamos *producto tensorial, directo o de Kronecker* de \mathcal{H}_1 por \mathcal{H}_2 a un espacio vectorial $\mathcal{H}_1 \otimes \mathcal{H}_2$ provisto de una aplicación bilineal de $\mathcal{H}_1 \times \mathcal{H}_2$ en $\mathcal{H}_1 \otimes \mathcal{H}_2$:

$$(\psi(1), \phi(2)) \rightarrow \psi(1) \otimes \phi(2) \equiv |\psi(1) \phi(2)\rangle$$

tal que, si $\{\dots\psi_i(1)\dots\}$ y $\{\dots\phi_j(2)\dots\}$ son bases de \mathcal{H}_1 y \mathcal{H}_2 respectivamente, los productos $\{\dots\psi_i(1) \otimes \phi_j(2)\dots\}$ forman una base de $\mathcal{H}_1 \otimes \mathcal{H}_2^*$. Consecuencias inmediatas de esta definición son:

— $\text{Dim}(\mathcal{H}_1 \otimes \mathcal{H}_2) = \text{Dim}(\mathcal{H}_1) \text{Dim}(\mathcal{H}_2)$,

— si $\psi(1) = \sum_i c_i \psi_i(1)$ y $\phi(2) = \sum_j d_j \phi_j(2)$,

$$\psi(1) \otimes \phi(2) = \sum_{ij} c_i d_j \psi_i(1) \phi_j(2),$$

pero existen elementos de $\mathcal{H}_1 \otimes \mathcal{H}_2$ no expresables como producto tensorial de un vector de \mathcal{H}_1 y otro de \mathcal{H}_2 ,

— si \mathcal{H}_1 y \mathcal{H}_2 están provistos de sendos productos escalares, se puede definir el siguiente producto escalar en $\mathcal{H}_1 \otimes \mathcal{H}_2$:

$$\langle \psi(1) \phi(2) | \psi'(1) \phi'(2) \rangle = \langle \psi(1) | \psi'(1) \rangle \langle \phi(2) | \phi'(2) \rangle.$$

Sean $\hat{A}(1)$ y $\hat{B}(2)$ dos *operadores* que actúan sobre los espacios vectoriales \mathcal{H}_1 y \mathcal{H}_2

* Se puede demostrar que el espacio $\mathcal{H}_1 \otimes \mathcal{H}_2$ existe y está determinado salvo isomorfismos.

respectivamente. Se define el producto tensorial $\hat{A}(1) \otimes \hat{B}(2)$ como un operador que actúa sobre $\mathcal{H}_1 \otimes \mathcal{H}_2$ y que cumple:

$$\{\hat{A}(1) \otimes \hat{B}(2)\} \psi(1) \otimes \phi(2) = \{\hat{A}(1) \psi(1)\} \otimes \{\hat{B}(2) \phi(2)\} .$$

Cualquier otro operador del espacio producto se puede expresar como combinación lineal de productos directos de operadores

$$\hat{C}(1, 2) = \sum_{ij} c_{ij} \hat{A}_i(1) \otimes \hat{B}_j(2) .$$

Los operadores

$$\hat{\hat{A}}(1) \equiv \hat{A}(1) \otimes \hat{1}(2)$$

y

$$\hat{\hat{B}}(2) \equiv \hat{1}(1) \otimes \hat{B}(2)$$

se conocen como extensiones de $\hat{A}(1)$ y $\hat{B}(2)$ a $\mathcal{H}_1 \otimes \mathcal{H}_2$, respectivamente. Cuando ello no induzca a confusión, prescindiremos de la tilde al referirnos a tales extensiones.

Ejercicio 1.26. Comprueba las siguientes afirmaciones:

a) $[\hat{A}(1), \hat{B}(2)] = 0$.

b) El proyector sobre $|\psi(1) \phi(2)\rangle$ es

$$|\psi(1) \phi(2)\rangle \langle \psi(1) \phi(2)| = |\psi(1)\rangle \langle \psi(1)| \otimes |\phi(2)\rangle \langle \phi(2)| .$$

c) Si $\{\dots \psi_{ij}(1) \dots\}$ es un conjunto completo \mathcal{H}_1 formado por vectores propios del operador autoadjunto $\hat{A}(1)$ y $\{\dots \phi_k(2) \dots\}$ es un conjunto completo de \mathcal{H}_2 , los vectores $\{\dots \psi_{ij}(1) \otimes \phi_k(2) \dots\}$, que forman un conjunto completo de $\mathcal{H}_1 \otimes \mathcal{H}_2$, son propios de $\hat{\hat{A}}(1)$; si a_i es un valor propio de $\hat{A}(1)$ de degeneración d_i en \mathcal{H}_1 , su degeneración en $\mathcal{H}_1 \otimes \mathcal{H}_2$ es $d_i \times \text{Dim}(\mathcal{H}_2)$; el proyector sobre a_i en $\mathcal{H}_1 \otimes \mathcal{H}_2$ es la extensión a este espacio del proyector sobre a_i en \mathcal{H}_1 .

d) Si $\hat{A}(1) \psi_i(1) = a_i \psi_i(1)$ y $\hat{B}(2) \phi_j(2) = b_j \phi_j(2)$,

$$\{\hat{A}(1) + \hat{B}(2)\} \psi_i(1) \otimes \phi_j(2) = (a_i + b_j) \psi_i(1) \otimes \phi_j(2)$$

(esto se puede generalizar a operadores con espectro continuo).

iii) Los subsistemas a los que hace referencia el postulado son, normalmente, partículas o conjuntos de partículas que pueden tratarse como partículas debido a que su estructura interna no se ve afectada por los experimentos que se van a considerar, como ocurre con los núcleos atómicos en la mayor parte de las reacciones químicas o con los átomos de helio cuando se estudia el comportamiento estadístico del elemento a bajas temperaturas. No obstante, la "partición" de un sistema puede entenderse también como partición de los observables de un CCOC en subconjuntos; por ejemplo, el espacio de Hilbert asociado a una partícula con spin se puede expresar como producto directo del espacio que le corresponde si se prescinde del spin

por el espacio de estados de spin (vease ejercicio 1.27). Esto permite utilizar los resultados del estudio de un sistema bajo unas condiciones determinadas como punto de partida para un estudio más profundo del mismo sistema que exija ampliar el CCOC.

En efecto, si $\{\dots\psi_{ij}\dots\}$ y $\{\dots\phi_{kl}\dots\}$ son conjuntos completos de \mathcal{H}_1 y de \mathcal{H}_2 , respectivamente, formados por vectores propios de los operadores autoadjuntos \hat{A} y \hat{B} , los productos $\{\dots\psi_{ij} \otimes \phi_{kl}\dots\}$ proporcionan automáticamente una base del espacio $\mathcal{H}_1 \otimes \mathcal{H}_2$ en cuyos elementos siguen estando bien definidos los observables A y B, (ejercicio 1.26. c)). En dichos estados el sistema compuesto es, a todos los efectos una superposición de subsistemas ya que cualquier valor esperado de una propiedad de uno de éstos calculada en el sistema compuesto toma el mismo valor que se obtendría estudiando el subsistema por separado:

$$\langle \psi_1(1) \otimes \dots \psi_i(i) \dots | \hat{1}(1) \otimes \dots \hat{A}_i(i) \otimes \dots | \psi_1(1) \otimes \dots \psi_i(i) \dots \rangle = \langle \psi_i(i) | \hat{A}_i(i) | \psi_i(i) \rangle.$$

Si, además, el hamiltoniano del sistema se puede expresar como suma de un término que solo contiene operadores del espacio \mathcal{H}_1 y otro que solo contiene operadores de \mathcal{H}_2 (en cuyo caso diremos que el hamiltoniano es separable), existirá un conjunto completo de estados estacionarios producto directo de un estado de \mathcal{H}_1 y otro de \mathcal{H}_2 (ejercicio 1.26. d)), forma que se mantendrá en los intervalos de tiempo de evolución libre del sistema. Es el caso de la caja de potencial tridimensional o del oscilador armónico tridimensional isótropo, cuyos estados estacionarios se pueden expresar como productos de estados estacionarios de tres sistemas unidimensionales asociadas a cada uno de los tres ejes cartesianos.

La discusión anterior explica la necesidad de recurrir al producto directo para estudiar sistemas compuestos: si un sistema está formado por dos partes que no interaccionan entre sí, podremos estudiarlas por separado y asignar estados "individuales" a cada una de ellas. Cada pareja de estados de una y otra parte será un estado del sistema compuesto y, si escogemos en cada parte un conjunto completo de estados y los combinamos de todas las formas posibles, tendremos un conjunto de parejas de vectores cuyas combinaciones lineales permitirán describir cualquier situación física del sistema compuesto, es decir, una base de su espacio de Hilbert. Si, como ocurre normalmente, la introducción de interacción entre las partes no modifica esencialmente la naturaleza del problema, aquellas parejas de vectores seguirán siendo una base del espacio de Hilbert asociado al sistema con interacción. No obstante, el hamiltoniano deja de ser separable y los estados estacionarios dejan de ser expresables como producto de estados de cada parte (salvo para formas muy especiales del término de interacción). Esta es la razón por la cual los productos de estados monoeléctricos no describen exactamente estados estacionarios de sistemas polielectricos y si lo pueden hacer, en cambio, combinaciones lineales de tales productos (interacción de configuraciones).

Ejercicio 1.27. El espacio de Hilbert de una partícula con spin se puede expresar como $\mathcal{H} = \mathcal{H}_o \otimes \mathcal{H}_s$, donde \mathcal{H}_o es el espacio que resulta al eliminar del CCOC el observable asociado al spin

y \mathcal{H}_s es el espacio de estados de spin. a) Si, para $\phi \in \mathcal{H}_o$ y $g \in \mathcal{H}_s$ se cumplen las relaciones $\hat{h}\phi = \epsilon\phi$, $\hat{l}_z\phi = m\phi$ y $\hat{s}_z g = m_s g$, indica si el vector $\phi \otimes g$ es propio de \hat{h} , de $\hat{l}_z \otimes \hat{s}_z$ y de $\hat{l} \otimes \hat{s}$ y, en caso de serlo, con qué valores propios. b) ¿Podremos expresar los estados estacionarios como producto de una parte espacial por una de spin si incluimos un término de interacción spin-órbita de la forma $\hat{l} \otimes \hat{s}$ en el hamiltoniano de la partícula? ¿y si el término fuera de la forma $\hat{l}_z \otimes \hat{s}_z$?

iv) Dos subsistemas se consideran idénticos cuando no pueden distinguirse mediante los observables relevantes en los experimentos a considerar. Por ejemplo dos núcleos de una misma molécula con iguales número atómico y número másico en sus respectivos estados internos de menor energía serán idénticos, a efectos químicos, si no es posible identificarlos mediante su posición en la molécula. Así, dos núcleos de hidrógeno de una molécula gaseosa de metano son indistinguibles y lo mismo ocurre con cualquier pareja de hidrógenos de un grupo metilo si la rotación del grupo respecto del resto de la molécula no está impedida. Sin embargo, los hidrógenos de la molécula CH_2ClBr son identificables mediante su posición en relación al plano que contiene a los núcleos C, Cl y Br, ya que no existe operación de simetría físicamente realizable por la molécula (rotación propia o cambio conformacional de barrera energética baja) que los intercambie.

Ya que los problemas de indistinguibilidad afectan, normalmente, a partículas o a conjuntos de partículas cuya estructura interna no sufre alteraciones relevantes, utilizaremos, en adelante, la denominación "partícula" para referirnos a cualquiera de tales subsistemas.

La posibilidad de identificar partículas suficientemente localizadas mediante su posición evita tener que preocuparse de que los vectores de estado estén convenientemente simetrizados respecto de intercambios de las partículas del sistema con todas las demás partículas del mismo tipo que existen en el universo, lo cual haría impracticable cualquier estudio cuántico. En sentido estricto, la posición nunca permitiría identificar partículas, ya que no es físicamente posible localizarlas con precisión absoluta en una región finita del espacio (las densidades de probabilidad siempre tienden a cero asintóticamente); no obstante, es fácil comprobar que, si las partículas están suficientemente localizadas, se obtiene la misma información física (mismos valores esperados) a partir de vectores de estado simetrizados que utilizando vectores sin simetrizar. El razonamiento anterior puede extenderse a observables distintos de la posición: si dos partículas idénticas en sus propiedades intrínsecas (carga, masa, módulo del spin, ...) difieren en una propiedad no intrínseca (por ejemplo, la componente del spin sobre un eje) y sólo consideramos experimentos en los que ésta no se ve afectada, podemos utilizar dicha propiedad para identificar las partículas, es decir, podemos considerarlas distinguibles. La propiedad utilizada puede estar bien definida en el sistema considerado o no estarlo y la

correspondiente distribución de probabilidades puede variar con el tiempo; en tal caso, basta que, en los experimentos e intervalos de tiempo a considerar, sea despreciable el solapamiento entre las densidades de probabilidad asociadas a una y otra partícula.

La definición de partículas idénticas dada al comienzo de este comentario exige los valores esperados de los observables relevantes del sistema sean invariantes frente al intercambio de aquéllas. Veamos como afecta dicho intercambio a los vectores y operadores que determinan tales valores esperados.

Dado un sistema de N partículas cuyo espacio de Hilbert asociado es $\mathcal{H} = \mathcal{H}_1 \otimes \dots \mathcal{H}_N$, el efecto del intercambio de las partículas 'i' y 'j' sobre un vector de la base $\{\dots, \psi_k(1) \otimes \dots \psi_l(i) \otimes \dots \psi_m(j) \otimes \dots \psi_n(N), \dots\}$ se puede establecer de dos maneras físicamente equivalentes: intercambiando los índices (i, j) que identifican las partículas o permutando los subíndices (l, m) que indican sus respectivos estados. Adoptaremos este último convenio, de forma que el intercambio produzca un vector del mismo espacio $\mathcal{H}_1 \otimes \dots \mathcal{H}_i \otimes \dots \mathcal{H}_j \otimes \dots \mathcal{H}_N$ (y no uno del espacio $\mathcal{H}_1 \otimes \dots \mathcal{H}_j \otimes \dots \mathcal{H}_i \otimes \dots \mathcal{H}_N$), y designaremos como $\widehat{\mathcal{P}}_{ij}$ el operador que efectúa el intercambio:

$$\widehat{\mathcal{P}}_{ij} \{\psi_k(1) \otimes \dots \psi_l(i) \otimes \dots \psi_m(j) \otimes \dots \psi_n(N)\} = \psi_k(1) \otimes \dots \psi_m(i) \otimes \dots \psi_l(j) \otimes \dots \psi_n(N).$$

Si las N partículas son idénticas, los espacios asociados a cada una de ellas también lo serán:

$\mathcal{H}_1 = \mathcal{H}_2 = \dots = \mathcal{H}_N$ y los identificaremos mediante un índice entre paréntesis: $\mathcal{H} = \mathcal{H}_1(1) \otimes \dots \mathcal{H}_1(N)$. La invarianza de los valores esperados frente a intercambios entre partículas idénticas impone cierta simetría sobre los correspondientes operadores. En efecto, dado un observable A se ha de cumplir

$$\begin{aligned} \langle \widehat{A} \rangle_{\Psi} &= \langle \Psi(1, 2, \dots, N) | \widehat{A}(1, 2, \dots, N) \Psi(1, 2, \dots, N) \rangle \\ &= \langle \widehat{\mathcal{P}}_{ij} \Psi(1, 2, \dots, N) | \widehat{A}(1, 2, \dots, N) \widehat{\mathcal{P}}_{ij} \Psi(1, 2, \dots, N) \rangle \\ &= \langle \Psi(1, 2, \dots, N) | \widehat{\mathcal{P}}_{ij}^{\dagger} \widehat{A}(1, 2, \dots, N) \widehat{\mathcal{P}}_{ij} \Psi(1, 2, \dots, N) \rangle \end{aligned}$$

para cualquier vector de estado Ψ , lo que implica

$$\widehat{A} = \widehat{\mathcal{P}}_{ij}^{\dagger} \widehat{A} \widehat{\mathcal{P}}_{ij}$$

y, como los operador $\widehat{\mathcal{P}}_{ij}$ son unitarios (ejercicio 1.28 a),

$$[\widehat{A}, \widehat{\mathcal{P}}_{ij}] = 0.$$

Esto se ha de cumplir para cualquier par de índices (i, j), y diremos que los operadores que representan observables del sistema han de ser simétricos respecto de los intercambios de partículas idénticas.

El razonamiento anterior se extiende fácilmente a permutaciones arbitrarias de las N partículas. En efecto, la extensión del convenio adoptado para intercambios de parejas de partículas a una permutación arbitraria $\{1, \dots, N\} \rightarrow \{\alpha_1, \dots, \alpha_N\}$ conduce a la siguiente definición del operador de permutación correspondiente:

$$\widehat{\mathcal{P}}_{\alpha} \{\psi_{i_1}(1) \otimes \dots \psi_{i_1}(N)\} = \psi_{i_{\alpha_1}}(1) \otimes \dots \psi_{i_{\alpha_N}}(N).$$

Teniendo en cuenta que toda permutación se puede expresar como producto de transposiciones (permutaciones de dos elementos), se demuestra fácilmente que los operadores $\widehat{\mathcal{P}}_\alpha$ son unitarios (ejercicio 1.28 b)), de manera que podemos apelar de nuevo a la invarianza de los valores esperados para concluir que los operadores que representan observables del sistema han de ser invariantes frente a permutaciones entre partículas idénticas:

$$[\widehat{A}, \widehat{\mathcal{P}}_\alpha] = 0.$$

Supongamos que el observable (o conjunto de observables) A es un CCOC del sistema y consideremos el proyector sobre un vector propio de \widehat{A} : $\widehat{P}_a \equiv |a\rangle\langle a|$. Este operador es también un observable del sistema (sería el observable asociado al aparato o conjunto de aparatos que permiten preparar el estado puro $|a\rangle$) y, por lo tanto, ha de conmutar con los operadores de permutación:

$$[\widehat{P}_a, \widehat{\mathcal{P}}_\alpha] = 0 \quad \forall \widehat{\mathcal{P}}_\alpha,$$

de donde se deduce que el vector $|a\rangle$ ha de ser propio de todos los $\widehat{\mathcal{P}}_\alpha$:

$$\widehat{\mathcal{P}}_\alpha (|a\rangle\langle a|) |\Psi\rangle = (|a\rangle\langle a|) \widehat{\mathcal{P}}_\alpha |\Psi\rangle,$$

es decir,

$$\widehat{\mathcal{P}}_\alpha |a\rangle = (\langle a | \widehat{\mathcal{P}}_\alpha | \Psi \rangle / \langle a | \Psi \rangle) |a\rangle.$$

Mediante teoría de grupos se demuestra que el grupo de las permutaciones de N elementos solo tiene dos representaciones irreducibles unidimensionales, a saber, la totalmente simétrica, de carácter $+1$ para todos los operadores de permutación, y la totalmente antisimétrica, de carácter $+1$ ó -1 según sea par o impar la permutación, es decir, según se descomponga en un número par o impar de transposiciones, número que designaremos π_α . Esto determina los posibles valores propios de los operadores $\widehat{\mathcal{P}}_\alpha$:

$$\widehat{\mathcal{P}}_\alpha |a\rangle = +1 |a\rangle \quad \forall \widehat{\mathcal{P}}_\alpha,$$

o bien

$$\widehat{\mathcal{P}}_\alpha |a\rangle = (-1)^{\pi_\alpha} |a\rangle \quad \forall \widehat{\mathcal{P}}_\alpha.$$

Los vectores simétricos constituyen un subespacio de $\mathcal{H}_1(1) \otimes \dots \mathcal{H}_1(N)$ que designaremos mediante la notación

$$\mathcal{H}_1^{\otimes_s N} \equiv \mathcal{H}_1(1) \otimes^s \dots \mathcal{H}_1(N).$$

Dada una base $\{\dots \psi_k \dots\}$ de \mathcal{H}_1 , los vectores de la forma

$$\begin{aligned} |(\psi_k \dots \psi_n)_+\rangle &\equiv \psi_k(1) \otimes^s \dots \psi_n(N) \\ &= (1/\sqrt{N!}) \sum_{\alpha=1}^{N!} \widehat{\mathcal{P}}_\alpha \{\psi_k(1) \otimes \dots \psi_n(N)\}, \end{aligned}$$

donde la suma se extiende a las permutaciones de N elementos, constituyen una base de aquel subespacio (vease ejercicio 1.18 d)). De forma análoga se introduce el subespacio antisimétrico:

$$\mathcal{H}_1^{\otimes_a N} \equiv \mathcal{H}_1(1) \otimes^a \dots \mathcal{H}_1(N)$$

una base del cual la forman los vectores:

$$\begin{aligned} |(\psi_k \dots \psi_n)_-\rangle &\equiv \psi_k(1) \otimes^a \dots \psi_n(N) \\ &= (1/\sqrt{N!}) \sum_{\alpha=1}^{N!} (-1)^{\pi_\alpha} \widehat{\mathcal{P}}_\alpha \{ \psi_k(1) \otimes \dots \psi_n(N) \} \end{aligned}$$

(vease ejercicio 1.18 d). Es evidente que los subespacios $\mathcal{H}_1^{\otimes sN}$ y $\mathcal{H}_1^{\otimes aN}$ son ortogonales entre sí, y se demuestra que las bases $\{ \dots | (\dots \psi_1 \dots \psi_m \dots)_\pm \rangle \dots \}$ son ortonormales si lo es $\{ \dots \psi_1 \dots \}$ *.

Dado que la paridad de toda permutación aumenta una unidad al componerla con una transposición, los vectores $| (\dots \psi_1 \dots \psi_m \dots)_\pm \rangle$ serán, respectivamente, simétricos o antisimétricos respecto del intercambio de dos partículas idénticas cualesquiera, de modo que, de acuerdo con el presente postulado, los vectores adecuados para representar sistemas de partículas idénticas con spin entero (bosones) o semiimpar (fermiones) serán elementos de $\mathcal{H}_1^{\otimes sN}$ y $\mathcal{H}_1^{\otimes aN}$, respectivamente. Se puede comprobar que tanto la evolución libre como el colapso del estado se producen dentro uno de aquellos subespacios, de manera que no existen contradicciones entre el presente postulado y los anteriores.

Ejercicio 1.28. Demuestra las siguientes afirmaciones:

- Los operadores de transposición $\widehat{\mathcal{P}}_{ij}$ son hermíticos y unitarios.
- Los operadores de permutación son unitarios.
- El *antisimetrizador* $\widehat{\mathcal{A}} \equiv \frac{1}{N!} \sum_{\alpha=1}^{N!} (-1)^{\pi_\alpha} \widehat{\mathcal{P}}_\alpha$ es un operador de proyección.
- Los vectores $| (\dots \psi_1 \dots \psi_m \dots)_+ \rangle$ (vease texto precedente) son simétricos frente a cualquier permutación de las partículas del sistema, y los vectores $| (\dots \psi_1 \dots \psi_m \dots)_- \rangle$ son simétricos respecto de las permutaciones pares y antisimétricos respecto de las impares.

1.6.1 Sistemas N-electrónicos

De acuerdo con el comentario iv) del sexto postulado, dada una base $\{ \dots \psi_i \dots \}$ del espacio de Hilbert asociado a un electrón, el conjunto de todos los vectores de la forma $| (\psi_{i_1} \dots \psi_{i_N})_- \rangle$ es una base del espacio asociado a un sistema de N electrones. Esta base es ortonormal si lo es aquélla. En efecto, sean $\Psi = | (\psi_1 \dots \psi_N)_- \rangle$ y $\Psi' = | (\psi'_1 \dots \psi'_N)_- \rangle$ dos productos directos antisimétricos con sus factores en máxima coincidencia (es decir, si existen estados monoeléctrónicos comunes a ambos productos han de ocupar las mismas posiciones en uno y otro),

$$\begin{aligned} \langle \Psi | \Psi' \rangle &= \langle \sqrt{N!} \widehat{\mathcal{A}} \{ \psi_1 \dots \psi_N \} | \sqrt{N!} \widehat{\mathcal{A}} \{ \psi'_1 \dots \psi'_N \} \rangle \\ &= \sum_{\alpha=1}^{N!} (-1)^{\pi_\alpha} \langle \psi_1 \dots \psi_N | \widehat{\mathcal{P}}_\alpha (\psi'_1 \dots \psi'_N) \rangle \\ &= \begin{cases} 1 & \text{si } \psi_i = \psi'_i \quad \forall i \\ 0 & \text{en caso contrario.} \end{cases} \end{aligned}$$

* En el apartado 1.6.1 lo demostraremos para el caso antisimétrico.

Consideremos dos operadores polielectrónicos \hat{F} y \hat{G} de la forma

$$\hat{F}(1, \dots, N) = \sum_{i=1}^N \hat{f}(i) ,$$

$$\hat{G}(1, \dots, N) = \sum_{i=1}^N \sum_{j>i}^N \hat{g}(i, j)$$

con $\hat{g}(i, j)$ simétrico respecto de la permutación de los electrones 'i' y 'j':

$$[\hat{g}(i, j), \hat{P}_{ij}] = 0$$

y un vector N-electrónico *de referencia* Ψ :

$$\Psi = |(\psi_1 \dots \psi_a \dots \psi_b \dots \psi_N)_-\rangle .$$

Llamaremos vectores monoelectrónicos *ocupados* a los que aparecen en el vector N-electrónico de referencia: $\psi_1, \dots, \psi_a, \dots, \psi_b, \dots, \psi_N$, y *virtuales* a los restantes: $\psi_{N+1}, \dots, \psi_r, \dots, \psi_s, \dots$. Un vector N-electrónico *monosustituido* será de la forma

$$\Psi_a^r = |(\psi_1 \dots \psi_a \dots \psi_b \dots \psi_N)_-\rangle ,$$

uno *disustituido*

$$\Psi_{ab}^{rs} = |(\psi_1 \dots \psi_r \dots \psi_s \dots \psi_N)_-\rangle ,$$

etc., y supondremos que todos ellos en máxima coincidencia con el de referencia. Con esta notación se cumplen las siguientes *relaciones de Condon-Slater*:

$$\langle \Psi | \hat{F} \Psi \rangle = \sum_{a=1}^N \langle \psi_a | \hat{f} \psi_a \rangle$$

$$\langle \Psi | \hat{F} \Psi_a^r \rangle = \langle \psi_a | \hat{f} \psi_r \rangle$$

$$\langle \Psi | \hat{F} \Psi_{ab}^{rs} \rangle = 0$$

...

$$\langle \Psi | \hat{G} \Psi \rangle = \sum_{a=1}^N \sum_{b>a}^N \langle \psi_a \psi_b || \psi_a \psi_b \rangle$$

$$\langle \Psi | \hat{G} \Psi_a^r \rangle = \sum_{b=1}^N \langle \psi_a \psi_b || \psi_r \psi_b \rangle$$

$$\langle \Psi | \hat{G} \Psi_{ab}^{rs} \rangle = \langle \psi_a \psi_b || \psi_r \psi_s \rangle$$

$$\langle \Psi | \hat{G} \Psi_{abc}^{rst} \rangle = 0$$

...

siendo

$$\langle \psi_a \psi_b || \psi_r \psi_s \rangle \equiv \langle \psi_a \psi_b | \psi_r \psi_s \rangle - \langle \psi_a \psi_b | \psi_s \psi_r \rangle .$$

Observese que esta expresión se anula cuando $\psi_r = \psi_s$ y es simétrica respecto del intercambio de los índices a-b y r-s simultáneamente; esto permite escribir el valor esperado de \hat{G} en la forma alternativa

$$\langle \Psi | \hat{G} \Psi \rangle = (1/2) \sum_{a=1}^N \sum_{b=1}^N \langle \psi_a \psi_b || \psi_a \psi_b \rangle .$$

Ejercicio 1.29. a) Demuestra que todo operador simétrico respecto de las permutaciones de electrones conmuta con el antisimetrizador.

b) Demuestra que todo operador del tipo \hat{F} es simétrico y que

$$\langle \Psi | \hat{F} \Psi \rangle = \sum_{a=1}^N \langle \psi_a | \hat{f} \psi_a \rangle .$$

c) Demuestra que todo operador del tipo \widehat{G} anterior es simétrico y que

$$\langle \Psi | \widehat{G} \Psi \rangle = \sum_{a=1}^N \sum_{b>a}^N \langle \psi_a \psi_b | | \psi_a \psi_b \rangle .$$

d) Expresa $\langle \Psi | \widehat{H} \Psi \rangle$ en función de estados monoeléctronicos para un sistema de N electrones cuyo hamiltoniano, en unidades atómicas, sea

$$\widehat{H} = \sum_i^N \widehat{h}(i) + \sum_{j>i}^N \widehat{r}_{ij}^{-1}$$

y comprueba que, si los estados monoeléctronicos son de la forma $\phi \otimes g$, (con la notación del ejercicio 1.27) donde g es una de las funciones α o β , propias del operador \widehat{s}_z con valores propios $\hbar/2$ y $-\hbar/2$ respectivamente,

$$\langle \Psi | \widehat{H} \Psi \rangle = \sum_{a=1}^{N_\alpha} h_{aa} + \sum_{a=1}^{N_\beta} h_{aa} + \sum_{a=1}^{N_\alpha} \sum_{b>a}^{N_\alpha} J_{ab} - K_{ab} + \sum_{a=1}^{N_\beta} \sum_{b>a}^{N_\beta} J_{ab} - K_{ab} + \sum_{a=1}^{N_\alpha} \sum_{b=1}^{N_\beta} J_{ab} ,$$

donde N_α y N_β son los números de estados monoeléctronicos con función de spin α y β , respectivamente, y la notación utilizada es la indicada en la sección 2.4. ¿Qué representación física puede asignarse a cada término de la expresión anterior?

2. REPRESENTACIONES

2.0 Introducción

De la misma forma que, para identificar un vector de \mathcal{R}^3 , hemos de referirnos a un sistema de coordenadas, como puede ser el asociado a la base ortonormal $\{\vec{i}, \vec{j}, \vec{k}\}$, deberemos recurrir también a una base para representar los vectores del espacio de Hilbert asociado a un sistema cuántico. La forma habitual de llevar a cabo esta *representación* consiste en elegir como base ortonormal las funciones propias de un CCOC que incluya observables significativos de cara al tipo de estudio que vayamos a realizar; por ejemplo, una base adecuada para representar los estados de un átomo de hidrógeno de cara a analizar su espectro será la formada por los vectores propios del CCOC $\{H, l^2, l_z \text{ y } s_z\}$ (suponemos que los efectos relativistas no son relevantes), ya que estos vectores representan estados en los que está bien definida la energía, que es la propiedad a medir, y los observables que determinan las reglas de selección. La expresión de un estado puro cualquiera en dicha base será:

$$\Psi = \sum_{nlmm_s} c_{nlmm_s} \psi_{nlmm_s}$$

$$\text{con } c_{nlmm_s} = \langle \psi_{nlmm_s} | \Psi \rangle$$

y la representación del estado en la base será el conjunto de coeficientes $\{\dots, c_{nlmm_s}, \dots\}$, que puede agruparse en una matriz columna \mathbf{c} tal como suele hacerse con las componentes $\{V_x, V_y, V_z\}$ de un vector de \mathcal{R}^3 en una base cartesiana.

2.1 Bases discretas

Dada una base discreta, no necesariamente ortonormal, $\{\psi_1, \dots, \psi_i, \dots\}$ del espacio de Hilbert, cualquier vector Ψ de este espacio se representa en aquella mediante una matriz columna \mathbf{c} cuyos elementos c_i son los coeficientes que expresan el vector como combinación lineal de los elementos de la base. Agrupando los elementos de ésta en una matriz fila:

$$\Psi \equiv (\psi_1, \dots, \psi_i, \dots),$$

podemos escribir

$$\Psi = \psi \mathbf{c}.$$

Un operador \hat{A} queda determinado sabiendo como actúa sobre los elementos de la base:

$$\hat{A} \psi_i = \sum_j \psi_j \mathcal{A}_{ji}, \quad i = 1, 2, \dots,$$

ya que, para cualquier vector Ψ del espacio,

$$\hat{A} \Psi = \hat{A} \sum_i \psi_i c_i = \sum_{ij} \psi_j \mathcal{A}_{ji} c_i.$$

Esta información puede expresarse en forma de matriz cuadrada \mathcal{A} con $(\mathcal{A})_{ji} = \mathcal{A}_{ji}$:

$$\boxed{\hat{A} \boldsymbol{\psi} = \boldsymbol{\psi} \mathcal{A}}$$
 ,

donde $\hat{A}(\psi_1, \dots, \psi_i, \dots) \equiv (\hat{A}\psi_1, \dots, \hat{A}\psi_i, \dots)$, y

$$\begin{aligned} \hat{A} \boldsymbol{\Psi} &= \hat{A} \boldsymbol{\psi} \mathbf{c} \\ &= \boldsymbol{\psi} \mathcal{A} \mathbf{c} . \end{aligned}$$

Nos referiremos a la matriz \mathcal{A} como *representación* del operador \hat{A} en la base $\{\psi_1, \dots, \psi_i, \dots\}$.

Si la base es ortonormal, los elementos de \mathcal{A} se calculan fácilmente multiplicando escalarmente la primera de estas ecuaciones por cada elemento de la base:

$$\langle \psi_j | \hat{A} \psi_i \rangle = \mathcal{A}_{ji} .$$

En cambio, si no lo es, la relación entre estos productos escalares, que designaremos con la notación A_{ji} , y la matriz \mathcal{A} pasa a ser

$$\begin{aligned} A_{ji} &\equiv \langle \psi_j | \hat{A} \psi_i \rangle \\ &= \sum_k \langle \psi_j | \psi_k \rangle \mathcal{A}_{ki} \end{aligned}$$

o, introduciendo la matriz de recubrimientos \mathbf{S} de elementos $S_{jk} \equiv \langle \psi_j | \psi_k \rangle$,

$$\boxed{\mathbf{A} = \mathbf{S} \mathcal{A}}$$
 .

Nos referiremos a la matriz \mathbf{A} como *representación según el producto escalar* del operador \hat{A} en la base $\{\psi_1, \dots, \psi_i, \dots\}$.

Se comprueba fácilmente que las matrices $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$ que representan los operadores $\hat{A}, \hat{B}, \hat{C}, \dots$ en una base se multiplican de la misma forma que éstos independientemente de que sea o no ortonormal la base, cosa que no ocurre con las representaciones según el producto escalar (ejercicio 2.1 b)). En este sentido, las primeras matrices son las auténticas representaciones de los correspondientes operadores, aunque las segundas juegan un papel muy destacado en la química cuántica.

Una consecuencia inmediata de lo anterior es que, si $\hat{B} = \hat{A}^{-1}$, $\mathcal{B} = \mathcal{A}^{-1}$ pero, en general, $\mathbf{B} \neq \mathbf{A}^{-1}$ e, incluso $\mathbf{A} \mathbf{B} \neq \mathbf{B} \mathbf{A}$. Sin embargo, si $\hat{B} = \hat{A}^\dagger$, $\mathbf{B} = \mathbf{A}^\dagger \equiv (\mathbf{A}^t)^*$ (matriz adjunta de \mathbf{A}), pero, en general, $\mathcal{B} \neq \mathcal{A}^\dagger$ (ejercicio 2.1 c)). En particular, la representación según el producto escalar de un operador autoadjunto en una base no ortonormal es una matriz hermítica (que coincide con su adjunta).

Observemos que la representación según el producto escalar del operador identidad es la matriz de recubrimientos de la base

$$(\mathbf{1})_{ij} = \langle \psi_i | \hat{1} \psi_j \rangle = S_{ij}$$

y se designará siempre como \mathbf{S} , reservando la notación $\mathbf{1}$ para la matriz identidad (que es la auténtica representación del operador identidad).

Ejercicio 2.1. a) Demuestra que, si $\Psi = \boldsymbol{\psi} \mathbf{c}$ donde $\boldsymbol{\psi}$ es una base no necesariamente ortonormal,

$$\langle \psi_k | \Psi \rangle = (\mathbf{S} \mathbf{c})_k .$$

b) Comprueba que, si $\widehat{A}\widehat{B} = \widehat{C}$, $\mathcal{A} \mathcal{B} = \mathbf{C}$, pero, en general, $\mathbf{A}\mathbf{B} \neq \mathbf{C}$.

c) Comprueba que, si $\widehat{B} = \widehat{A}^\dagger$, $\mathbf{B} = \mathbf{A}^\dagger$ pero, en general, $\mathcal{B} \neq \mathcal{A}^\dagger$.

d) Demuestra que

$$\widehat{A} = \sum_{ijk} |\psi_i\rangle \mathcal{A}_{ij} (\mathbf{S}^{-1})_{jk} \langle \psi_k| ,$$

donde \mathcal{A} es la representación de \widehat{A} en la base discreta $\boldsymbol{\psi}$. Indica la forma que tomará el desarrollo anterior cuando la base es ortonormal y cuando, además, sus vectores son propios del operador \widehat{A} . Utiliza la expresión anterior para obtener la resolución de la identidad en una base discreta no ortonormal:

$$\widehat{1} = \sum_{ij} |\psi_i\rangle (\mathbf{S}^{-1})_{ij} \langle \psi_j| .$$

e) Comprueba que $(\mathbf{A}\mathbf{B})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$

2.1.1 Cambios de base

Un cambio de base $\{\dots \psi_i \dots\} \rightarrow \{\dots \psi'_r \dots\}$ puede caracterizarse mediante la matriz cuadrada \mathbf{L} definida mediante

$$\psi'_r = \sum_i \psi_i L_{ir}$$

ó

$$\boldsymbol{\psi}' = \boldsymbol{\psi} \mathbf{L}$$

y afecta de la siguiente forma a las representaciones matriciales de los vectores:

$$\Psi = \boldsymbol{\psi} \mathbf{c} = \boldsymbol{\psi}' \mathbf{c}' = \boldsymbol{\psi} \mathbf{L} \mathbf{c}'$$

de donde

$$\boldsymbol{\psi}(\mathbf{c} - \mathbf{L} \mathbf{c}') = 0$$

y, como los elementos de $\boldsymbol{\psi}$ son linealmente independientes,

$$\mathbf{c} = \mathbf{L} \mathbf{c}'$$

y, si la matriz \mathbf{L} tiene inversa (lo que sucederá siempre que ambas bases definan el mismo espacio),

$$\mathbf{c}' = \mathbf{L}^{-1} \mathbf{c} .$$

Análogamente, dado un operador \widehat{A} ,

$$\widehat{A} \boldsymbol{\psi} = \boldsymbol{\psi} \mathcal{A}$$

y

$$\hat{A} \boldsymbol{\psi}' = \boldsymbol{\psi}' \boldsymbol{\mathcal{A}}';$$

multiplicando la primera relación por la derecha por \mathbf{L} y utilizando la segunda se obtiene

$$\boldsymbol{\psi}(\boldsymbol{\mathcal{A}} \mathbf{L} - \mathbf{L} \boldsymbol{\mathcal{A}}') = 0$$

y, utilizando de nuevo la independencia lineal de los vectores de la base,

$$\boxed{\boldsymbol{\mathcal{A}} \mathbf{L} = \mathbf{L} \boldsymbol{\mathcal{A}}'}$$

Si existe \mathbf{L}^{-1} ,

$$\boldsymbol{\mathcal{A}} = \mathbf{L} \boldsymbol{\mathcal{A}}' \mathbf{L}^{-1}$$

o

$$\boldsymbol{\mathcal{A}}' = \mathbf{L}^{-1} \boldsymbol{\mathcal{A}} \mathbf{L}.$$

La ley de transformación para las representaciones según el producto escalar es (ejercicio 2.2. b))

$$\boxed{\mathbf{A}' = \mathbf{L}^\dagger \mathbf{A} \mathbf{L}}$$

de donde se deduce que cualquier transformación \mathbf{L} que cumpla

$$\mathbf{L}^\dagger \mathbf{S} \mathbf{L} = \mathbf{1}$$

transforma la base $\{\dots \psi_i \dots\}$ en una base ortonormal $\{\dots \psi'_r \dots\}$.

Ejercicio 2.2. Demuestra que, si $\boldsymbol{\psi}' = \boldsymbol{\psi} \mathbf{L}$,

a) $\mathbf{S}' = \mathbf{L}^\dagger \mathbf{S} \mathbf{L}$;

b) para cualquier operador \hat{A} ,

$$\mathbf{A}' = \mathbf{L}^\dagger \mathbf{A} \mathbf{L};$$

c) si $\boldsymbol{\psi}$ y $\boldsymbol{\psi}'$ son ortonormales, $\mathbf{L}^{-1} = \mathbf{L}^\dagger$ (es decir, la matriz \mathbf{L} es unitaria).

Ejercicio 2.3. Demuestra que la traza de un operador es igual a la traza de su representación en cualquier base (sea o no ortonormal). ¿Podemos decir lo mismo de su representación según el producto escalar?

2.1.2 Valores y vectores propios

La expresión matricial de la ecuación de *valores propios* de un operador autoadjunto se obtiene de forma inmediata a partir de lo anterior. En efecto, si $\{\dots \psi_i \dots\}$ es un conjunto ortonormal completo de vectores propios de \hat{A} ($\hat{A} \psi_i = a_i \psi_i$), la representación de este operador en aquella base será diagonal:

$$\hat{A} \boldsymbol{\psi} = \boldsymbol{\psi} \mathbf{a}_\psi$$

con $(\mathbf{a}_\psi)_{ij} = a_i \delta_{ij}$

y viceversa. Conocida la representación \mathcal{A}_ϕ de \hat{A} en una base cualquiera $\{\dots \phi_i \dots\}$, el problema de valores propios de \hat{A} se reduce a hallar el cambio de base $\boldsymbol{\psi} = \boldsymbol{\phi} \mathbf{L}$ que diagonaliza la matriz \mathcal{A}_ϕ :

$$\mathcal{A}_\phi \mathbf{L} = \mathbf{L} \mathbf{a}_\psi .$$

como se comprueba introduciendo este cambio de base en la primera ecuación:

$$\hat{A} \boldsymbol{\phi} \mathbf{L} = \boldsymbol{\phi} \mathbf{L} \mathbf{a}_\psi$$

i teniendo en cuenta que $\hat{A} \boldsymbol{\phi} = \boldsymbol{\phi} \mathcal{A}_\phi$.

Llamando \mathbf{l}_i a la columna i-ésima de \mathbf{L} es inmediato comprobar que la ecuación $\mathcal{A}_\phi \mathbf{l}_i = \mathbf{l}_i a_i$ equivale al conjunto de ecuaciones:

$$\mathcal{A}_\phi \mathbf{l}_i = a_i \mathbf{l}_i$$

con $i = 1, 2, \dots$

Este conjunto es lo que se conoce como ecuación de valores propios de la matriz \mathcal{A}_ϕ , y es equivalente a la ecuación de valores propios del operador ($\hat{A} \psi_i = a_i \psi_i$), como se comprueba tomando $\psi_i = \boldsymbol{\psi} \mathbf{l}_i$. Por lo tanto, cada vector columna \mathbf{l}_i contiene los coeficientes de un vector propio de \hat{A} con valor propio a_i .

Multiplicando por la izquierda por \mathbf{S}_ϕ la ecuación de diagonalización de \mathcal{A}_ϕ podemos expresarla en función de representaciones según el producto escalar:

$$\mathbf{A}_\phi \mathbf{L} = \mathbf{S}_\phi \mathbf{L} \mathbf{a}_\psi .$$

En el ejercicio 2.1. e) vimos que los operadores hermíticos se representan según el producto escalar mediante matrices hermíticas, para las cuales existen algoritmos de diagonalización eficientes. Por esta razón, conviene utilizar bases ortonormales en las cuales la ecuación anterior expresa la diagonalización de la matriz \mathbf{A}_ϕ :

$$\mathbf{A}_\phi \mathbf{L} = \mathbf{L} \mathbf{a}_\psi$$

o, teniendo en cuenta el resultado del ejercicio 2.2.c),

$$\mathbf{a}_\psi = \mathbf{L}^\dagger \mathbf{A}_\phi \mathbf{L} .$$

Ejercicio 2.4. a) Comprueba que una matriz real ortogonal ($\mathbf{L}^t = \mathbf{L}^{-1}$) 2×2 solo tiene un parámetro independiente y que se puede expresar de la siguiente forma

$$\mathbf{L} = \begin{pmatrix} \cos\alpha & -\text{sen}\alpha \\ \text{sen}\alpha & \cos\alpha \end{pmatrix}$$

con $\alpha \in \mathcal{R}$.

b) Demuestra que una matriz \mathbf{A} real simétrica 2×2 se puede diagonalizar mediante una transformación del tipo indicado en a) con

$$\text{tg}2\alpha = 2 A_{12} / (A_{11} - A_{22}) .$$

El método de Jacobi para diagonalizar matrices reales simétricas es un procedimiento iterativo basado en la solución analítica planteada en el ejercicio anterior para el caso bidimensional. Cada iteración consiste en seleccionar el elemento no diagonal a_{ij} de mayor valor absoluto y efectuar una "rotación 2x2" que transforme en cero aquel elemento (y su simétrico) y deje inalteradas las filas y columnas de índices distintos de 'i' y 'j'. Las iteraciones se suceden hasta que todos los elementos no diagonales tienen valores absolutos inferiores a un cierto umbral. Aunque este método es muy sencillo se han desarrollado otros mucho más eficientes numéricamente que permiten diagonalizar matrices de gran tamaño o, por lo menos, determinar algunos de sus valores y vectores propios.

La diagonalización de matrices proporciona una vía práctica para determinar la matriz que representa a una función analítica de un operador

$$f(\hat{A}) = \sum_{k=0}^{\infty} f_k \hat{A}^k$$

conocida la representación matricial de éste en una base ortonormal discreta, $\{\dots \phi_i \dots\}$. En efecto,

$$\begin{aligned} \langle \phi_i | f(\hat{A}) | \phi_j \rangle &= \sum_{k=0}^{\infty} f_k \langle \phi_i | \hat{A}^k | \phi_j \rangle \\ &= \sum_{k=0}^{\infty} f_k \sum_l \sum_m \dots \sum_p \langle \phi_i | \hat{A} | \phi_l \rangle \langle \phi_l | \hat{A} | \phi_m \rangle \dots \langle \phi_p | \hat{A} | \phi_j \rangle \\ &= \sum_{k=0}^{\infty} f_k (\mathbf{A}^k)_{ij} \\ &= [f(\mathbf{A})]_{ij}, \end{aligned}$$

donde se ha definido la función f de la matriz \mathbf{A} como

$$f(\mathbf{A}) \equiv \sum_{k=0}^{\infty} f_k \mathbf{A}^k.$$

Esta función puede calcularse fácilmente previa diagonalización de \mathbf{A} . En efecto, sea

$$\mathbf{A} = \mathbf{L} \mathbf{a} \mathbf{L}^{-1}$$

con \mathbf{a} diagonal,

$$\begin{aligned} f(\mathbf{A}) &= \sum_{k=0}^{\infty} f_k \mathbf{L} \mathbf{a} \mathbf{L}^{-1} \dots \mathbf{L} \mathbf{a} \mathbf{L}^{-1} \\ &= \mathbf{L} \left(\sum_{k=0}^{\infty} f_k \mathbf{a}^k \right) \mathbf{L}^{-1} \\ &= \mathbf{L} f(\mathbf{a}) \mathbf{L}^{-1}, \end{aligned}$$

donde $f(\mathbf{a})$ se calcula de manera trivial (ejercicio 2.5. a)).

Ejercicio 2.5. a) Demuestra que, si \mathbf{a} es una matriz diagonal, $(f(\mathbf{a}))_{ij} = f(a_i) \delta_{ij}$.

b) Demuestra que, si $f(x)$ es una función real analítica de la variable real x , $[f(\mathbf{A})]^\dagger = f(\mathbf{A}^\dagger)$ (ten en cuenta el apartado e) del ejercicio 2.1). En particular, si \mathbf{A} es hermítica, también lo será $f(\mathbf{A})$.

2.1.3 Ortonormalización

En el apartado anterior hemos podido apreciar el interés que presenta la utilización de bases ortonormales en relación con el cálculo de los valores y vectores propios de un operador auto-adjunto. Este es uno de los muchos problemas que se presentan en química cuántica en los que la ortogonalidad de la base simplifica considerablemente el aparato matemático, por lo cual interesa disponer de métodos para *ortonormalizar* conjuntos discretos de vectores linealmente independientes, es decir, transformarlos en conjuntos ortonormales que engendren el mismo espacio. Esta transformación no es única y, de hecho distintos métodos conducirán, en general a distintas bases ortonormales.

Una forma sencilla de efectuar dicha ortonormalización es el método propuesto por *Schmidt*: si $\{\psi_1, \dots, \psi_i, \dots\}$ es el conjunto no ortonormal de partida, puede obtenerse un nuevo conjunto ortonormal $\{\psi_1', \dots, \psi_i', \dots\}$ del siguiente modo:

$$\psi_1' = \psi_1 / \langle \psi_1 | \psi_1 \rangle^{1/2},$$

$$c_2 \psi_2' = \psi_2 - \langle \psi_1' | \psi_2 \rangle \psi_1'$$

$$\text{con } c_2 \text{ tal que } \langle \psi_2' | \psi_2' \rangle = 1,$$

$$c_3 \psi_3' = \psi_3 - \langle \psi_1' | \psi_3 \rangle \psi_1' - \langle \psi_2' | \psi_3 \rangle \psi_2'$$

$$\text{con } c_3 \text{ tal que } \langle \psi_3' | \psi_3' \rangle = 1, \text{ ect.,}$$

es decir, restamos de cada vector su proyección sobre el subespacio engendrado por los anteriores vectores ortonormalizados y lo multiplicamos por una constante de normalización.

Este método exige elegir previamente una ordenación para los vectores a ortonormalizar y esta elección influye sobre el resultado de la ortonormalización, ya que los últimos vectores sufrirán, en general, una modificación mayor que los primeros. Una alternativa que no adolece de este inconveniente es el método de ortonormalización *simétrica*, propuesto por Löwdin, que consiste en aplicar al conjunto inicial de vectores la transformación de matriz $\mathbf{L} = \mathbf{S}^{-1/2}$, donde \mathbf{S} es la matriz de recubrimientos de aquel conjunto (vease ejercicio 2.6).

Ejercicio 2.6. Demuestra que la transformación $\mathbf{L} = \mathbf{S}^{-1/2}$ aplicada a un conjunto de vectores linealmente independientes cuya matriz de recubrimientos es \mathbf{S} conduce a un conjunto ortonormal e indica como calcularías $\mathbf{S}^{-1/2}$.

La matriz $\mathbf{S}^{-1/2}$ no puede calcularse si \mathbf{S} tiene algún valor propio nulo, lo cual indicaría que el correspondiente vector propio es combinación lineal de los restantes y, por lo tanto, el conjunto de vectores de partida sería redundante. Esto no ocurrirá en una base bien elegida, pero puede suceder que el cálculo de $\mathbf{S}^{-1/2}$ presente problemas numéricos debido a la existencia de valores propios muy pequeños en \mathbf{S} , en cuyo caso el método de ortonormalización simétrica no

será aplicable. Una forma de solventar este problema consiste en utilizar el método de ortonormalización *canónica*, que se basa en la utilización de la matriz de transformación

$$\mathbf{L} = \mathbf{U} \mathbf{s}^{-1/2},$$

donde $\mathbf{U}^\dagger \mathbf{S} \mathbf{U} = \mathbf{s}$ y \mathbf{s} es diagonal. Si \mathbf{S} tiene n valores propios de los cuales k son inferiores a un cierto umbral, podemos eliminar los correspondientes vectores propios y construir un conjunto ortonormal de $n-k$ elementos utilizando la matriz de transformación rectangular

$$\overline{\mathbf{L}} \equiv \overline{\mathbf{U}} \overline{\mathbf{s}}^{-1/2},$$

donde $\overline{\mathbf{s}}^{-1/2}$ es la matriz diagonal que contiene los $n-k$ mayores valores propios de \mathbf{S} y $\overline{\mathbf{U}}$ es la matriz rectangular cuyas columnas contienen los correspondientes vectores propios. Este método de ortonormalización permite, pues, reducir el tamaño de una base que presenta cuasi-dependencias lineales alterando lo mínimo posible el espacio engendrado por aquélla.

2.2 Bases truncadas

Dado un conjunto $\{\chi_1, \dots, \chi_n\}$ no completo de vectores linealmente independientes de un espacio vectorial \mathcal{H} o "base truncada", la matriz $\overline{\mathcal{A}}_\chi$ definida mediante

$$\widehat{\mathbf{A}} \overline{\boldsymbol{\chi}} = \overline{\boldsymbol{\chi}} \overline{\mathcal{A}}_\chi$$

con

$$\overline{\boldsymbol{\chi}} = (\chi_1 \dots \chi_n)$$

no es una auténtica representación de $\widehat{\mathbf{A}}$, ya que tales matrices no se multiplican, en general, como los correspondientes operadores. Aquella matriz, convenientemente completada con ceros, es, en cambio, la representación de la "proyección del operador $\widehat{\mathbf{A}}$ sobre el subespacio \mathcal{H}_χ^- " engendrado por los vectores $\{\chi_1, \dots, \chi_n\}$, operador que definiremos como

$$\widehat{\mathbf{P}}_\chi^- \widehat{\mathbf{A}} \widehat{\mathbf{P}}_\chi^- ,$$

donde $\widehat{\mathbf{P}}_\chi^-$ es el proyector sobre dicho subespacio. En efecto, completemos la base truncada con una base del complemento ortogonal de \mathcal{H}_χ^- en \mathcal{H} : $\{\chi_{n+1}, \dots\}$. El proyector sobre \mathcal{H}_χ^- actuará de la siguiente forma sobre la base $\{\chi_1, \dots, \chi_n, \chi_{n+1}, \dots\}$ de \mathcal{H} :

$$\widehat{\mathbf{P}}_\chi^- \chi_j = \chi_j \quad \text{si } 1 \leq j \leq n ,$$

$$\widehat{\mathbf{P}}_\chi^- \chi_j = 0 \quad \text{si } j > n ,$$

de modo que, para $1 \leq j \leq n$,

$$(\widehat{\mathbf{P}}_\chi^- \widehat{\mathbf{A}} \widehat{\mathbf{P}}_\chi^-) \chi_j = \widehat{\mathbf{P}}_\chi^- \widehat{\mathbf{A}} \chi_j = \widehat{\mathbf{P}}_\chi^- \sum_{i=1}^n \chi_i (\mathcal{A}_\chi^-)_{ij} = \sum_{i=1}^n \chi_i (\mathcal{A}_\chi^-)_{ij}$$

y, para $j > n$,

$$(\widehat{\mathbf{P}}_\chi^- \widehat{\mathbf{A}} \widehat{\mathbf{P}}_\chi^-) \chi_j = 0 .$$

La representación del operador $\widehat{\mathbf{P}}_\chi^- \widehat{\mathbf{A}} \widehat{\mathbf{P}}_\chi^-$ en la base $\{\chi_1, \dots, \chi_n, \chi_{n+1}, \dots\}$ será, pues, una matriz de dimensiones iguales a la de \mathcal{H} cuyo primer cuadrante $n \times n$ coincide con la matriz $\overline{\mathcal{A}}_\chi$ y con

todos los demás elementos iguales a cero:

$$\widehat{P}_{\chi}^{-} \widehat{A} \widehat{P}_{\chi}^{-} (\chi_1, \dots, \chi_n, \chi_{n+i}, \dots) = (\chi_1, \dots, \chi_n, \chi_{n+i}, \dots) \begin{pmatrix} \mathcal{A}_{\chi}^{-} & 0 \\ 0 & 0 \end{pmatrix}.$$

Ejercicio 2.7. a) Teniendo en cuenta el enunciado del ejercicio 2.1 (apartado d), comprueba que

$$\widehat{P}_{\chi}^{-} = \sum_{ij=1}^n |\chi_i\rangle (\overline{\mathbf{S}}^{-1})_{ij} \langle \chi_j|$$

es el proyector sobre el espacio engendrado por el conjunto de vectores linealmente independientes $\{\chi_1 \dots \chi_n\}$.

b) Obtén la representación del operador $\widehat{P}_{\chi}^{-} \widehat{A} \widehat{P}_{\chi}^{-}$ en una base $\{\chi_1, \dots, \chi_n, \chi_{n+1}, \dots\}$, obtenida completando el conjunto anterior con una base del complemento ortogonal del subespacio engendrado por aquel, utilizando la expresión anterior para \widehat{P}_{χ}^{-} y comprueba que se obtiene el resultado hallado en el texto.

c) Indica la forma que tendrá la representación según el producto escalar del operador $\widehat{P}_{\chi}^{-} \widehat{A} \widehat{P}_{\chi}^{-}$ en la base $\{\chi_1, \dots, \chi_n, \chi_{n+1}, \dots\}$.

La hermiticidad del operador $\widehat{P}_{\chi}^{-} \widehat{A} \widehat{P}_{\chi}^{-}$ nos garantiza que los valores propios de \mathcal{A}_{χ}^{-} sean reales y que vectores propios correspondientes a distinto valor propio sean ortogonales entre sí. Ahora bien, ¿existe alguna relación entre estos vectores y valores propios y los de \widehat{A} ? Para responder a esta pregunta observemos que cualquier valor propio de $\widehat{P}_{\chi}^{-} \widehat{A} \widehat{P}_{\chi}^{-}$ se puede expresar como valor esperado de \widehat{A} sobre el correspondiente vector propio:

$$\widehat{P}_{\chi}^{-} \widehat{A} \widehat{P}_{\chi}^{-} \psi = \bar{a} \psi$$

de donde

$$\begin{aligned} \bar{a} &= \langle \psi | \widehat{P}_{\chi}^{-} \widehat{A} \widehat{P}_{\chi}^{-} \psi \rangle \\ &= \langle \widehat{P}_{\chi}^{-} \psi | \widehat{A} \widehat{P}_{\chi}^{-} \psi \rangle \\ &= \langle \psi | \widehat{A} \psi \rangle \end{aligned}$$

y, de acuerdo con el *teorema variacional*, este valor esperado ha de ser superior al menor valor propio de \widehat{A} . Deduiremos este resultado para un operador de espectro discreto desarrollando ψ en un conjunto completo $\{\dots \phi_i \dots\}$ de vectores propios del operador:

$$\begin{aligned} \langle \psi | \widehat{A} \psi \rangle &= \sum_{ij} c_i^* \langle \phi_i | \widehat{A} \phi_j \rangle c_j \\ &= \sum_{ij} c_i^* a_i \langle \phi_i | \phi_j \rangle c_j \\ &= \sum_i |c_i|^2 a_i \\ &\geq \sum_i |c_i|^2 a_1 = a_1, \end{aligned}$$

donde el signo igual solo es válido si $\psi = \phi_1$, vector propio de \widehat{A} con menor valor propio a_1 .

Ejercicio 2.8. Deduce el teorema variacional utilizando únicamente el significado físico del

valor esperado.

Podríamos pensar en optimizar ψ como aproximación a ϕ_1 minimizando $\langle \psi | \hat{A} \psi \rangle$ respecto de los coeficientes l_i de la combinación lineal que lo expresa en la base truncada,

$$\psi = \sum_{i=1}^n \chi_i l_i = \bar{\chi} \mathbf{l},$$

pero vamos a ver que el vector óptimo es, precisamente, un vector propio de $\hat{P}_{\bar{\chi}} \hat{A} \hat{P}_{\bar{\chi}}$.

Sea $\langle A \rangle_{\psi}$ el valor esperado de \hat{A} en el estado ψ , que supondremos no necesariamente normalizado,

$$\langle A \rangle_{\psi} \equiv \frac{\langle \psi | \hat{A} \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\sum_{ij=1}^n l_i^* \langle \chi_i | \hat{A} \chi_j \rangle l_j}{\sum_{ij=1}^n l_i^* \langle \chi_i | \chi_j \rangle l_j} = \frac{\mathbf{l}^\dagger \mathbf{A}_{\bar{\chi}} \mathbf{l}}{\mathbf{l}^\dagger \bar{\mathbf{S}} \mathbf{l}}.$$

La variación producida en $\langle A \rangle_{\psi}$ por un cambio $\delta\psi = \sum_{i=1}^n \chi_i \delta l_i = \bar{\chi} \delta \mathbf{l}$ en ψ será

$$\begin{aligned} \delta \langle A \rangle_{\psi} &= \frac{(\delta \mathbf{l}^\dagger \mathbf{A}_{\bar{\chi}} \mathbf{l} + \mathbf{l}^\dagger \mathbf{A}_{\bar{\chi}} \delta \mathbf{l}) \mathbf{l}^\dagger \bar{\mathbf{S}} \mathbf{l} - \mathbf{l}^\dagger \mathbf{A}_{\bar{\chi}} \mathbf{l} (\delta \mathbf{l}^\dagger \bar{\mathbf{S}} \mathbf{l} + \mathbf{l}^\dagger \bar{\mathbf{S}} \delta \mathbf{l})}{(\mathbf{l}^\dagger \bar{\mathbf{S}} \mathbf{l})^2} = \\ &= \frac{\delta \mathbf{l}^\dagger \mathbf{A}_{\bar{\chi}} \mathbf{l} - \langle A \rangle_{\psi} \delta \mathbf{l}^\dagger \bar{\mathbf{S}} \mathbf{l}}{\mathbf{l}^\dagger \bar{\mathbf{S}} \mathbf{l}} + \text{complejo conjugado} = \\ &= \frac{\sum_{i=1}^n \delta l_i^* (\mathbf{A}_{\bar{\chi}} \mathbf{l} - \langle A \rangle_{\psi} \bar{\mathbf{S}} \mathbf{l})_i + \delta l_i (\mathbf{A}_{\bar{\chi}} \mathbf{l} - \langle A \rangle_{\psi} \bar{\mathbf{S}} \mathbf{l})_i^*}{\mathbf{l}^\dagger \bar{\mathbf{S}} \mathbf{l}}. \end{aligned}$$

Para que $\langle A \rangle_{\psi}$ sea mínimo, la expresión anterior deberá anularse para cualquier $\delta \mathbf{l}$ y, en particular, para variaciones con todas las componentes δl_i nulas excepto una cualquiera, que puede tomarse real o imaginaria pura. Por lo tanto, se han de anular todos los coeficientes que multiplican a los δl_i , es decir,

$$\mathbf{A}_{\bar{\chi}} \mathbf{l} = \langle A \rangle_{\psi} \bar{\mathbf{S}} \mathbf{l},$$

que es la ecuación de valores propios de la matriz $\mathcal{A}_{\bar{\chi}}$ expresada en función de $\mathbf{A}_{\bar{\chi}}$. Las soluciones no triviales de esta ecuación se obtienen para valores de $\langle A \rangle_{\psi}$ que anulen el determinante secular:

$$|\mathbf{A}_{\bar{\chi}} - \langle A \rangle_{\psi} \bar{\mathbf{S}}| = 0.$$

Las n raíces reales de esta ecuación se pueden reunir en una matriz diagonal $\bar{\mathbf{a}}$, lo que permite agrupar las correspondientes ecuaciones de valores propios en la ecuación matricial

$$\mathbf{A}_{\bar{\chi}} \mathbf{L} = \bar{\mathbf{S}} \mathbf{L} \bar{\mathbf{a}},$$

donde \mathbf{L} es la matriz $n \times n$ cuyas columnas son las matrices \mathbf{l} .

Los puntos estacionarios de $\langle A \rangle_{\psi}$ son, pues vectores propios de $\mathcal{A}_{\bar{\chi}}$ (o de $\hat{P}_{\bar{\chi}} \hat{A} \hat{P}_{\bar{\chi}}$) y el menor de éstos ha de corresponder al mínimo buscado. Si vamos completando el conjunto de vectores $\{\dots \chi_i \dots\}$, los valores y vectores propios de $\mathcal{A}_{\bar{\chi}}$ se irán aproximando a los exactos y, de acuerdo con el *teorema de separación*, cada adición de un nuevo vector al conjunto rebaja cada valor propio y añade uno nuevo que queda por encima de todos los anteriores. Por lo tanto, si ordenamos los valores propios de $\mathbf{A}_{\bar{\chi}}$ de menor a mayor, el i -ésimo será siempre una

cota superior del i -ésimo valor propio exacto de \hat{A} .

Dado el importante papel que juegan los estados estacionarios en la descripción de sistemas microscópicos, la aplicación más interesante de lo anterior se obtiene cuando \hat{A} es el hamiltoniano del sistema. Para otros observables puede tener poco interés el cálculo aproximado de sus valores y vectores propios si no conocemos la relación entre éstos últimos y los estados estacionarios del sistema.

Aunque la mayor parte de los cálculos que se realizan en química cuántica utilizan bases truncadas, nos referiremos con frecuencia a tales conjuntos sin explicitar su carácter incompleto, lo cual no debe hacernos olvidar que los resultados que produzcan tales cálculos no corresponderán a los observables exactos, sino a sus proyecciones sobre los subespacios engendrados por las bases truncadas.

2.3 Bases continuas

Debido a la importancia que tienen los observables posición y momento, interesa, con frecuencia, incluir unos u otros en el CCOC utilizado para construir la representación, de forma que las componentes de un vector de estado en la correspondiente base nos proporcionen información directa sobre la distribución de probabilidad asociada a unos u otros observables; se obtienen así las representaciones de *posiciones* y de *momentos*, respectivamente. Dado que estos observables son de espectro continuo, los coeficientes que representen un vector de estado en una de tales representaciones se identificarán mediante uno o más índices continuos; la función que expresa aquellos coeficientes para cada conjunto de valores de estos índices se conoce como *función de onda* del estado en la representación considerada. De acuerdo con la notación introducida en el ejercicio 1.6, las funciones de onda son amplitudes de densidad de probabilidad de los observables continuos que definen la representación.

Consideremos el caso más sencillo: una partícula unidimensional cuyos únicos observables tienen análogo clásico (por lo tanto no tendrá spin). Es fácil comprobar que tanto la posición como el momento constituyen, por si solos, un CCOC del sistema. En efecto, como todos los observables tienen análogo clásico, podrán expresarse como funciones de x y p . Por otra parte, el conmutador de \hat{x} y el operador asociado a un observable cualquiera $A(x, p)$ será (ejercicio 1.24. b))

$$[\hat{x}, \hat{A}] = i\hbar \frac{\partial \hat{A}}{\partial \hat{p}}$$

y se anulará si y solo si A no depende de p . Es decir, todo observable compatible con la posición será función únicamente de ésta, de modo que $\{x\}$ será un conjunto de observables compatibles maximal y, evidentemente, no redundante. Análogo razonamiento puede hacerse

con el momento.

Ejercicio 2.9. Generaliza el razonamiento anterior a sistemas de N partículas tridimensionales sin spin.

La función de onda en la representación de posiciones será

$$\Psi(x) = \langle \psi_x | \Psi \rangle \equiv \langle x | \Psi \rangle$$

y, en la de momentos,

$$\tilde{\Psi}(p) = \langle \psi_p | \Psi \rangle \equiv \langle p | \Psi \rangle.$$

De acuerdo con el tercer postulado (ejercicio 1.6), $|\Psi(x)|^2$ y $|\tilde{\Psi}(p)|^2$ representarán las densidades de probabilidad de obtener el valor 'x' y el valor 'p', respectivamente, al medir la posición y el momento de la partícula en el estado Ψ .

De la misma forma que, en \mathcal{R}^3 , el producto escalar entre dos vectores se puede expresar como suma de los productos de sus componentes en una base cartesiana, también podemos calcular el producto escalar entre dos vectores de estado de una partícula unidimensional mediante la integral de los productos de sus respectivas funciones de onda:

$$\langle \Psi | \Psi' \rangle = \langle \Psi | \int_{-\infty}^{\infty} |x\rangle \langle x| dx | \Psi' \rangle = \int_{-\infty}^{\infty} \Psi^*(x) \Psi'(x) dx = \int_{-\infty}^{\infty} \tilde{\Psi}^*(p) \tilde{\Psi}'(p) dp.$$

Ejercicio 2.10. Indica como se expresará el producto escalar en la representación asociada al CCOC $\{x, y, z, s_z\}$ de un sistema monoeléctrico.

La representación de posiciones (lo mismo que la de momentos) de un operador será una "matriz" de índices continuos, es decir, una función de dos variables:

$$A(x, x') = \langle \psi_x | \hat{A} \psi_{x'} \rangle = \langle x | \hat{A} | x' \rangle.$$

Para el operador posición dicha representación es, evidentemente, diagonal:

$$\langle x | \hat{x} | x' \rangle = x \delta(x'-x),$$

de donde se deduce que su efecto sobre $\Psi(x)$ se reduce a multiplicar esta función por la variable 'x':

$$\langle x | \hat{x} \Psi \rangle \equiv (\hat{x}\Psi)(x) = \langle x | \hat{x} \left(\int_{-\infty}^{\infty} |x'\rangle \langle x'| dx' \right) | \Psi \rangle = x \Psi(x).$$

La forma que toma un operador \hat{A} en una representación se puede designar indicando la representación entre paréntesis, con lo que la relación anterior se podrá escribir en la siguiente forma:

$$\hat{x}(x) \Psi(x) = x \Psi(x).$$

La forma usual de elegir el operador momento en la representación de posiciones es

$$\hat{p}(x) \Psi(x) = (-i\hbar d/dx) \Psi(x) .$$

La indicación de la representación utilizada se suele omitir si ello no induce a confusión.

Ejercicio 2.11. a) Comprueba que la forma adoptada para el operador \hat{p} en la representación de posiciones es acorde con el segundo postulado.

b) Indica si es correcta la siguiente definición alternativa para el operador \hat{p} en la representación de posiciones: $\hat{p}(x) = (-i\hbar d/dx) + \hbar df/dx$, siendo f una función real, analítica y adimensional de la variable x .

La representación de posiciones de un vector propio del operador momento se obtiene resolviendo la ecuación diferencial

$$\hat{p}(x) \psi_p(x) = p \psi_p(x) ,$$

cuyas soluciones son de la forma

$$\psi_p(x) = (2\pi\hbar)^{-1/2} \exp(ipx/\hbar) ,$$

donde la "constante de normalización" $(2\pi\hbar)^{-1/2}$ se introduce para que el producto escalar entre dos vectores propios del operador momento coincida con una de las expresiones que definen la delta de Dirac:

$$\langle p | p' \rangle = (2\pi\hbar)^{-1} \int_{-\infty}^{\infty} \exp[i(p'-p)x/\hbar] dx = \delta(p-p') .$$

Para la representación de momentos podríamos proceder de forma análoga, y el paso de una a otra se efectúa introduciendo el oportuno desarrollo de la identidad en el producto escalar que define la función de onda:

$$\tilde{\Psi}(p) = \langle p | \left(\int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) | \Psi \rangle = \int_{-\infty}^{\infty} \langle p | x \rangle \langle x | \Psi \rangle dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \exp(-ipx/\hbar) \Psi(x) dx$$

y

$$\Psi(x) = \int_{-\infty}^{\infty} \langle x | p \rangle \langle p | \Psi \rangle dp = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \exp(ipx/\hbar) \tilde{\Psi}(p) dp ,$$

donde se reconoce a $\tilde{\Psi}(p)$ como transformada de Fourier de $\Psi(x)$.

Ejercicio 2.12. a) Escribe, como función explícita del tiempo, la función de onda en la representación de momentos que corresponde a una partícula libre sin spin cuya función de onda en representación de posiciones es el "paquete de ondas" gaussiano (en unidades atómicas):

$$\Psi(x, t) = \int_{-\infty}^{\infty} g(p) e^{-iE_p t} \frac{1}{\sqrt{2\pi}} e^{ipx} dp$$

con $E_p = p^2/2m$ y $g(p) = \sqrt{\frac{2a^2}{\pi}} e^{-a^2(p-p_0)^2}$ ('a' y 'p₀' son constantes reales).

¿Cuál es la densidad de probabilidad $\sigma(p, t)$ correspondiente? ¿Representa un estado estacionario?

b) ¿En qué casos cabe esperar que el cuadrado del módulo de la función de onda en representación de posiciones ($|\Psi(x, t)|^2$) sea independiente del tiempo?

c) ¿Cabe esperar que varíen Δx y Δp con el tiempo en dicho estado?

d) Al calcular Δx a partir de $\Psi(x, t)$ se obtiene un valor que aumenta con el tiempo:

$$\Delta x = a \{1 + (t^2/4m^2a^4)\}^{1/2}$$

Razona físicamente el aumento de Δx con t . ¿Existe contradicción entre el resultado anterior y la reversidad temporal inherente a la ecuación de Schrödinger?

e) Calcula Δp y comprueba que, en el instante $t=0$, el producto de las dispersiones de los observables x y p es el mínimo permitido por la relación de dispersión.

$$\Delta x \Delta p = 1/2 .$$

La generalización de las relaciones anteriores a partículas tridimensionales se efectúa sin dificultad adoptando la siguiente normalización para los estados propios del momento en representación de posiciones:

$$\psi_{\vec{p}}(\vec{r}) = (2\pi\hbar)^{-3/2} \exp(i\vec{p}\cdot\vec{r}/\hbar) .$$

Observemos que los productos directos entre vectores de estado, necesarios, por ejemplo, para describir un sistema de varias partículas a partir de la descripción de cada una de ellas, se reducen a productos de funciones complejas cuando pasamos a una representación continua, y las propiedades de aquel producto directo se traducen a propiedades triviales del producto de funciones cuya imagen es el mismo cuerpo, los números complejos.

2.4 Notación

A continuación presentamos una compilación de la notación que utilizaremos para tratar sistemas N-electrónicos en representación de posiciones.

$\vec{x}_i \equiv i$: coordenadas de posición espaciales (\vec{r}_i) y de spin ($\omega_i \equiv m_{s_i}$) del electrón 'i'.

$\psi_j(\vec{x}_i) = \langle \vec{x}_i | \psi_j \rangle$: j-ésimo spin-orbital (vector de estado monoeléctrico en la representación

de posiciones) en las coordenadas del electrón 'i'; en general, se utilizan spin-orbitales de la forma

$$\psi_j(\vec{x}_i) = \phi_j(\vec{r}_i) \alpha(\omega_i)$$

o

$$\psi_j(\vec{x}_i) = \phi_j(\vec{r}_i) \beta(\omega_i),$$

donde

$$\phi_j(\vec{r}_i) = \langle \vec{r}_i | \phi_j \rangle$$

es el j-ésimo orbital en las coordenadas del electrón 'i' y

$$\alpha(\omega_i) = \langle m_{si} | \alpha \rangle = \delta_{m_{si}, 1/2},$$

$$\beta(\omega_i) = \langle m_{si} | \beta \rangle = \delta_{m_{si}, -1/2},$$

son las representaciones de posiciones de los estados propios de \hat{s}_z con valores propios 1/2 ($|\alpha\rangle$) y -1/2 ($|\beta\rangle$), respectivamente, en unidades atómicas.

$\chi_\mu(\vec{r}_i)$: μ -ésimo orbital de base en las coordenadas del electrón 'i'; la expresión de un orbital en la base $\chi = (\chi_1, \dots, \chi_n)$ será

$$\phi_j(\vec{r}_i) = \sum_{\mu=1}^n \chi_\mu(\vec{r}_i) c_{\mu i}.$$

$\Psi(\vec{x}_1, \dots, \vec{x}_N) = \langle \vec{x}_1 \dots \vec{x}_N | \Psi \rangle$: función de onda N-electrónica.

$$\begin{aligned} |\psi_1 \dots \psi_N| &= \langle \vec{x}_1 \dots \vec{x}_N | (\psi_1 \dots \psi_N) \rangle \\ &= \sqrt{N!} \hat{\mathcal{A}} (\psi_1(\vec{x}_1) \dots \psi_N(\vec{x}_N)) \\ &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\vec{x}_1) & \dots & \psi_N(\vec{x}_1) \\ \dots & \dots & \dots \\ \psi_1(\vec{x}_N) & \dots & \psi_N(\vec{x}_N) \end{vmatrix} : \end{aligned}$$

determinante de Slater N-electrónico. Cuando no se indique lo contrario, se supondrá que los determinantes de Slater están contruidos a partir de spin-orbitales ortonormales.

Ψ^0 ó Ψ : determinante de Slater de referencia; los subíndices a, b, ... se utilizarán para referir se a (spin) orbitales ocupados en Ψ^0 y los subíndices r, s, ... para (spin) orbitales no ocupados o virtuales.

$$\langle \psi_i | \hat{f} \psi_j \rangle = \sum_{\omega=-1/2}^{1/2} \int_{\mathcal{R}^3} \psi_i^*(\vec{x}) \hat{f}(\vec{x}) \psi_j(\vec{x}) d\vec{r} :$$

integral monoeléctronica (entre spin-orbitales); es habitual utilizar la notación simplificada

$$\langle \psi_i | \hat{f} \psi_j \rangle = \int \psi_i^*(\vec{x}) \hat{f}(\vec{x}) \psi_j(\vec{x}) d\vec{x}$$

o, lo que es lo mismo, tratar la variable ω_i como continua con $\alpha(\omega_i) = \delta(\omega_i-1/2)$ y $\beta(\omega_i) = \delta(\omega_i+1/2)$.

$$\langle \phi_i | \hat{h} | \phi_j \rangle = \int_{\mathcal{R}^3} \phi_i^*(\vec{r}) \hat{h}(\vec{r}) \phi_j(\vec{r}) d\vec{r} :$$

integral monoeléctrica (entre orbitales).

$$\begin{aligned} \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_k \psi_l \rangle &= \sum_{\omega_1, \omega_2 = -1/2}^{1/2} \int_{\mathcal{R}^6} \psi_i^*(\vec{x}_1) \psi_j^*(\vec{x}_2) \frac{1}{r_{12}} \psi_k(\vec{x}_1) \psi_l(\vec{x}_2) d\vec{r}_1 d\vec{r}_2 \\ &= \int \psi_i^*(\vec{x}_1) \psi_j^*(\vec{x}_2) \frac{1}{r_{12}} \psi_k(\vec{x}_1) \psi_l(\vec{x}_2) d\vec{x}_1 d\vec{x}_2 : \end{aligned}$$

integral bielectrónica (entre spin-orbitales).

$$\langle \phi_i \phi_j | \frac{1}{r_{12}} | \phi_k \phi_l \rangle = \int_{\mathcal{R}^6} \phi_i^*(\vec{r}_1) \phi_j^*(\vec{r}_2) \frac{1}{r_{12}} \phi_k(\vec{r}_1) \phi_l(\vec{r}_2) d\vec{r}_1 d\vec{r}_2 \equiv (\phi_i \phi_k | \phi_j \phi_l) :$$

integral bielectrónica (entre orbitales).

$J_{ij} = (\phi_i \phi_i | \phi_j \phi_j)$: integral de Coulomb,

$K_{ij} = (\phi_i \phi_j | \phi_j \phi_i)$: integral de intercambio.

3. MATRICES DE DENSIDAD REDUCIDAS

3.0 Introducción

En el comentario ii) del sexto postulado se indicó que los estados estacionarios de un sistema compuesto por dos partes que no interactúan entre sí se pueden representar mediante vectores de estado producto directo de vectores de cada parte:

$$\Psi(1, 2) = \psi(1) \otimes \phi(2) .$$

En un estado de este tipo podemos afirmar que el subsistema 1 está en el estado $\psi(1)$ y el subsistema 2 está en el estado $\phi(2)$, ya que, para un observable del subsistema 1,

$$\begin{aligned} \langle A(1) \rangle_{\Psi(1, 2)} &= \langle \psi(1) \otimes \phi(2) | \hat{A}(1, 2) | \psi(1) \otimes \phi(2) \rangle \\ &= \langle \psi(1) | \hat{A}(1) | \psi(1) \rangle \\ &= \langle A(1) \rangle_{\psi(1)} , \end{aligned}$$

y análogamente para uno del subsistema 2.

Esta asignación de vectores de estado particulares para cada subsistema no es posible cuando existe interacción entre las partes, pues los estados estacionarios serán de la forma

$$\Psi(1, 2) = \sum_{ij} c_{ij} \psi_i(1) \otimes \phi_j(2) ,$$

no expresable como producto directo de vectores de cada parte. El valor esperado de un observable del subsistema 1 en aquel estado será

$$\begin{aligned} \langle A(1) \rangle_{\Psi(1, 2)} &= \sum_{ijkl} c_{ij}^* c_{kl} \langle \psi_i(1) \otimes \phi_j(2) | \hat{A}(1, 2) | \psi_k(1) \otimes \phi_l(2) \rangle \\ &= \sum_{ik} \langle \psi_i(1) | \hat{A}(1) | \psi_k(1) \rangle \sum_j c_{ij}^* c_{kj} , \end{aligned}$$

y, en general, no se puede poner esta expresión en la forma $\langle A(1) \rangle_{\psi(1)}$ (a no ser que los coeficientes c_{ij} se anulen para todo $j \neq j_0$, en cuyo caso $\Psi(1, 2) = \{\sum_i c_{ij} \psi_i(1)\} \otimes \phi_{j_0}(2)$). Aunque no es posible asignar vectores de estado particulares a cada subsistema, podemos definir sendos operadores de densidad reducidos que determinan completamente las propiedades de aquéllos, como veremos a continuación.

3.1 Operadores de densidad reducidos

Sea $\hat{\rho}(1, 2)$ el operador densidad que describe un estado (puro o mezcla) del sistema compuesto 1-2, y sean $\{\dots \psi_i(1) \dots\}$ y $\{\dots \phi_i(2) \dots\}$ bases discretas ortonormales de los espacios de Hilbert asociados a los subsistemas 1 y 2,

$$\langle A(1) \rangle_{\rho} = \text{Tr} (\hat{A} \hat{\rho})$$

$$\begin{aligned}
 &= \sum_{ijkl} \langle \psi_i \phi_j | \hat{A}(1) \otimes \hat{I}(2) | \psi_k \phi_l \rangle \langle \psi_k \phi_l | \hat{\rho}(1, 2) | \psi_i \phi_j \rangle \\
 &= \sum_{ik} \langle \psi_i | \hat{A}(1) | \psi_k \rangle \sum_j \langle \psi_k \phi_j | \hat{\rho}(1, 2) | \psi_i \phi_j \rangle \\
 &= \text{Tr} [\hat{A}(1) \hat{\rho}_1(1)] ,
 \end{aligned}$$

donde hemos introducido el *operador de densidad reducido del subsistema 1* a través de su representación matricial en la base $\{\dots \psi_i(1) \dots\}$:

$$\langle \psi_k | \hat{\rho}_1(1) | \psi_i \rangle = \sum_j \langle \psi_k \phi_j | \hat{\rho}(1, 2) | \psi_i \phi_j \rangle = \langle \psi_k | \sum_j \langle \phi_j | \hat{\rho}(1, 2) | \phi_j \rangle | \psi_i \rangle .$$

La operación que produce $\hat{\rho}_1(1)$ a partir de $\hat{\rho}(1, 2)$ se conoce como *traza parcial respecto del subsistema 2* y se representa así:

$$\boxed{\hat{\rho}_1(1) = \sum_j \langle \phi_j | \hat{\rho}(1, 2) | \phi_j \rangle \equiv \text{Tr}_2 \hat{\rho}(1, 2)} .$$

También se dice que la matriz ρ_1 que representa al operador $\hat{\rho}_1$ en la base $\{\dots \psi_i \dots\}$ es la traza parcial respecto de la(-s) variable(-s) '2' de la matriz ρ que representa al operador $\hat{\rho}$ en la base $\{\dots \psi_i(1) \otimes \phi_j(2) \dots\}$:

$$\rho_1 = \text{Tr}_2 \rho .$$

Análogamente,

$$\hat{\rho}_2(2) = \text{Tr}_1 \hat{\rho}(1, 2)$$

y

$$\rho_2 = \text{Tr}_1 \rho .$$

En un estado puro de la forma $\Psi(1, 2) = \psi(1) \otimes \phi(2)$ ambos subsistemas se encuentran en estados puros descritos por los operadores de densidad $\hat{\rho}_1(1) = |\psi(1)\rangle \langle \psi(1)|$ y $\hat{\rho}_2(2) = |\phi(2)\rangle \langle \phi(2)|$ (ejercicio 3.1 b)). Además, $\hat{\rho}(1, 2) = \hat{\rho}_1(1) \otimes \hat{\rho}_2(2)$ y los estados de los subsistemas no están correlacionados. No obstante, existen estados puros del sistema compuesto que no pueden expresarse como producto de estados de cada subsistema, existiendo una correlación entre los estados de éstos. Es el caso, por ejemplo, de los dos electrones de un átomo de helio en su estado fundamental o, de un modo más general, de un estado de spin singulete de dos partículas de spin 1/2, independientemente de si existe o no interacción entre ellas (ejercicio 3.1 c)).

Ejercicio 3.1. a) Demuestra que $\text{Tr} \hat{\rho}(1, 2) = \text{Tr} (\text{Tr}_2 \hat{\rho}(1, 2)) = \text{Tr} (\text{Tr}_1 \hat{\rho}(1, 2))$.

b) Comprueba que, para un sistema en el estado puro descrito por el vector

$$\Psi(1, 2) = \psi(1) \otimes \phi(2)$$

con ψ y ϕ normalizados, se cumplen las siguientes igualdades:

$$\hat{\rho}_1(1) = |\psi(1)\rangle \langle \psi(1)| ,$$

$$\hat{\rho}_2(2) = |\phi(2)\rangle \langle \phi(2)|$$

y

$$\hat{\rho}(1, 2) = \hat{\rho}_1(1) \otimes \hat{\rho}_2(2) .$$

c) Dos partículas de spin 1/2 se encuentran en el estado de spin singulete descrito por la función

de onda $\{\alpha(1)\beta(2)-\beta(1)\alpha(2)\}/\sqrt{2}$. Comprueba que los operadores de densidad reducidos de cada partícula representan estados mezcla con 50% de α y 50% de β , y que su producto no coincide con el operador de densidad del sistema completo.

Las definiciones anteriores se generalizan de forma inmediata a sistemas con un número arbitrario N de partes:

$$\hat{\rho}_1(1) = \text{Tr}_{2\dots N} \hat{\rho}(1, 2, \dots, N)$$

con

$$\langle \psi_k | \hat{\rho}_1(1) | \psi_i \rangle = \sum_{j_2 \dots j_N} \langle \psi_k \phi_{j_2} \dots \phi_{j_N} | \hat{\rho}(1, 2, \dots, N) | \psi_i \phi_{j_2} \dots \phi_{j_N} \rangle .$$

Por otra parte, pueden utilizarse representaciones en bases continuas o mixtas (i. e., con índices discretos y continuos) para calcular la traza parcial, lo cual implica la oportuna substitución de sumatorios por integrales. En el siguiente apartado veremos ejemplos de representaciones de este tipo.

3.2 Matrices de densidad reducidas de sistemas polielectrónicos en representación de posiciones

3.2.1 Matriz de densidad reducida de orden 1

Los operadores de densidad reducidos proporcionan una vía muy conveniente para expresar las propiedades de los sistemas N -electrónicos. Se acostumbra expresar dichos operadores en la representación de posiciones, mediante matrices de índices continuos; por ejemplo, la matriz que representa el operador de densidad reducido asociado al electrón 1 será:

$$\begin{aligned} \rho_1(\vec{x}_1; \vec{x}_1') &\equiv \langle \vec{x}_1 | \hat{\rho}_1(1) | \vec{x}_1' \rangle \\ &= \langle \vec{x}_1 | \text{Tr}_{2\dots N} \hat{\rho}(1, 2, \dots, N) | \vec{x}_1' \rangle \\ &= \int \langle \vec{x}_1, \vec{x}_2, \dots, \vec{x}_N | \hat{\rho}(1, 2, \dots, N) | \vec{x}_1', \vec{x}_2, \dots, \vec{x}_N \rangle d\vec{x}_2 \dots d\vec{x}_N \end{aligned}$$

y, si el sistema se encuentra en un estado puro descrito por la función de onda $\Psi(1, 2, \dots, N)$,

$$\hat{\rho}(1, 2, \dots, N) = |\Psi(1, 2, \dots, N)\rangle \langle \Psi(1, 2, \dots, N)|$$

y

$$\rho_1(\vec{x}_1; \vec{x}_1') = \int \Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \Psi^*(\vec{x}_1', \vec{x}_2, \dots, \vec{x}_N) d\vec{x}_2 \dots d\vec{x}_N .$$

Como los electrones son indistinguibles, $\rho_1(\vec{x}; \vec{x}') = \rho_2(\vec{x}; \vec{x}') = \dots = \rho_N(\vec{x}; \vec{x}')$, y conviene definir la *matriz de densidad reducida de orden 1* como

$$\begin{aligned} \gamma(\vec{x}; \vec{x}') &\equiv \sum_{i=1}^N \rho_i(\vec{x}; \vec{x}') \\ &= N \rho_1(\vec{x}; \vec{x}') \end{aligned}$$

$$= N \int \Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \Psi^*(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) d\vec{x}_2 \dots d\vec{x}_N .$$

Ejercicio 3.2. Calcula la traza de la matriz de densidad reducida de orden 1.

El elemento "diagonal" $\gamma(\vec{x}; \vec{x})$ representa la densidad de probabilidad $\sigma(\vec{x})$ de encontrar un electrón cualquiera en el punto \vec{r} y con un valor ω para la coordenada de spin (independientemente de los valores que tomen las coordenadas de los restantes electrones). En efecto, la densidad de probabilidad correspondiente al electrón 1 sería:

$$\begin{aligned} \sigma_1(\vec{x}) &= \int |\Psi(\vec{x}, \vec{x}_2, \vec{x}_3, \dots, \vec{x}_N)|^2 d\vec{x}_2 d\vec{x}_3 \dots d\vec{x}_N \\ &= \rho_1(\vec{x}; \vec{x}) \end{aligned}$$

y, como los N electrones son indistinguibles,

$$\begin{aligned} \sigma(\vec{x}) &= \sum_{i=1}^N \sigma_i(\vec{x}) \\ &= N \sigma_1(\vec{x}) \\ &= N \rho_1(\vec{x}; \vec{x}) \\ &= \gamma(\vec{x}; \vec{x}) . \end{aligned}$$

Dado un operador del tipo $\hat{F}(1, \dots, N) = \sum_{i=1}^N \hat{f}(i)$,

$$\begin{aligned} \langle \hat{F} \rangle_{\Psi} &= N \langle \Psi | \hat{f}(1) | \Psi \rangle \\ &= N \text{Tr}(\hat{f} \hat{\rho}_1) \end{aligned}$$

e, introduciendo el operador $\hat{\gamma} = N \hat{\rho}_1$,

$$\boxed{\langle \hat{F} \rangle_{\Psi} = \text{Tr}(\hat{f} \hat{\gamma})} .$$

Si el operador \hat{f} tiene representación diagonal en posiciones,

$$\langle \vec{x} | \hat{f} | \vec{x}' \rangle = f(x) \delta(\vec{x} - \vec{x}') ,$$

la expresión anterior se reduce a

$$\begin{aligned} \langle \hat{F} \rangle_{\Psi} &= \int \langle \vec{x} | \hat{f} | \vec{x}' \rangle \langle \vec{x}' | \hat{\gamma} | \vec{x} \rangle d\vec{x} d\vec{x}' \\ &= \int f(\vec{x}) \gamma(\vec{x}; \vec{x}) d\vec{x} . \end{aligned}$$

Ciertos operadores monoelectrónicos, si bien no son diagonales en la representación de posiciones, son "locales" en el sentido de que su efecto sobre una función $\psi(\vec{x})$ en un punto \vec{x}_0 se puede determinar conociendo el valor de esta función en un pequeño entorno del punto*. Este es el caso del operador momento:

* Operadores no locales son aquéllos cuyo efecto sobre una función en un punto depende de los valores que toma la función en una región finita; se llaman también operadores integrales por razones obvias:

$$\langle \vec{x} | \hat{A} | \psi \rangle = \int \langle \vec{x} | \hat{A} | \vec{x}' \rangle \langle \vec{x}' | \psi \rangle d\vec{x}' = \int A(\vec{x}; \vec{x}') \psi(\vec{x}') d\vec{x}' .$$

$$\begin{aligned}\langle x | \hat{p} \psi \rangle &= (\hat{p}\psi)(x) \\ &= (-i\hbar d/dx) \psi(x) \\ &= \hat{p}(x) \langle x | \psi \rangle.\end{aligned}$$

Para operadores de este tipo,

$$\begin{aligned}\langle F \rangle_{\Psi} &= \int \langle \vec{x} | \hat{f} \hat{\gamma} | \vec{x} \rangle d\vec{x} \\ &= \int \{ \langle \vec{x}' | \hat{f} \hat{\gamma} | \vec{x} \rangle \}_{\vec{x}'=\vec{x}} d\vec{x}\end{aligned}$$

y, teniendo en cuenta que $\langle \vec{x}' | \hat{f}(1) \psi \rangle = \hat{f}(\vec{x}') \langle \vec{x}' | \psi \rangle$,

$$\begin{aligned}\langle F \rangle_{\Psi} &= \int \{ f(\vec{x}') \langle \vec{x}' | \hat{\gamma} | \vec{x} \rangle \}_{\vec{x}'=\vec{x}} d\vec{x} \\ &= \int \{ f(\vec{x}') \gamma(\vec{x}'; \vec{x}) \}_{\vec{x}'=\vec{x}} d\vec{x},\end{aligned}$$

expresión que generaliza la obtenida para operadores diagonales.

Con frecuencia encontraremos operadores locales que no dependen de las coordenadas de spin, lo cual hace conveniente definir la *matriz de densidad reducida de orden 1 sin spin* como la traza parcial respecto de la coordenada de spin de la matriz de densidad de orden 1:

$$\begin{aligned}P(\vec{r}; \vec{r}') &\equiv \text{Tr}_{\omega} \gamma(\vec{x}; \vec{x}') \\ &= \int \gamma(\vec{r} \omega; \vec{r}' \omega) d\omega.\end{aligned}$$

Con esta definición, el valor esperado de un operador que, en representación de posiciones, es de la forma

$$\hat{F}(\vec{x}_1, \dots, \vec{x}_N) = \sum_{i=1}^N \hat{f}(\vec{r}_i)$$

con $\hat{f}(\vec{r}_i)$ local se puede expresar como

$$\begin{aligned}\langle F \rangle_{\Psi} &= \int_{\vec{r}} \{ f(\vec{r}') \int_{\omega} \gamma(\vec{r}' \omega; \vec{r} \omega) d\omega \}_{\vec{r}'=\vec{r}} d\vec{r} \\ &= \int \{ f(\vec{r}') P(\vec{r}'; \vec{r}) \}_{\vec{r}'=\vec{r}} d\vec{r}\end{aligned}$$

o, introduciendo el operador \hat{P} cuya representación en posiciones es $P(\vec{r}; \vec{r}')$,

$$\boxed{\langle F \rangle_{\Psi} = \text{Tr}(\hat{f} \hat{P})}.$$

Ejercicio 3.3. Comprueba que, para un estado puro Ψ ,

a) $P(\vec{r}; \vec{r}') = N \int \Psi(\vec{r} \omega, \vec{x}_2, \dots, \vec{x}_N) \Psi^*(\vec{r}' \omega, \vec{x}_2, \dots, \vec{x}_N) d\omega d\vec{x}_2 \dots d\vec{x}_N$,

b) $-\sigma(\vec{r}) \equiv -P(\vec{r}; \vec{r})$ representa la densidad de carga electrónica medida en unidades atómicas.

3.2.2 Matriz de densidad reducida de orden 2

Si estamos interesados en propiedades bielectrónicas, deberemos utilizar operadores de

densidad reducidos para subsistemas de dos electrones, como es

$$\hat{\rho}_{12}(1, 2) = \text{Tr}_{3\dots N} \hat{\rho}(1, 2, \dots N)$$

cuya representación en posiciones será

$$\begin{aligned} \rho_{12}(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2') &\equiv \langle \vec{x}_1, \vec{x}_2 | \hat{\rho}_{12}(1, 2) | \vec{x}_1', \vec{x}_2' \rangle = \\ &= \int \langle \vec{x}_1, \vec{x}_2, \vec{x}_3, \dots \vec{x}_N | \hat{\rho}(1, \dots N) | \vec{x}_1', \vec{x}_2', \vec{x}_3, \dots \vec{x}_N \rangle d\vec{x}_3 \dots d\vec{x}_N \end{aligned}$$

y, para un estado puro $\Psi(1, 2, \dots N)$,

$$\rho_{12}(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2') = \int \Psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots \vec{x}_N) \Psi^*(\vec{x}_1', \vec{x}_2', \vec{x}_3, \dots \vec{x}_N) d\vec{x}_3 \dots d\vec{x}_N .$$

A partir de estas representaciones se define la *matriz de densidad reducida de orden 2* como

$$\begin{aligned} \Gamma(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2') &\equiv N(N-1) \rho_{12}(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2') \\ &= N(N-1) \int \Psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots \vec{x}_N) \Psi^*(\vec{x}_1', \vec{x}_2', \vec{x}_3, \dots \vec{x}_N) d\vec{x}_3 \dots d\vec{x}_N , \end{aligned}$$

donde el factor $N(N-1)$ se introduce para que los elementos diagonales representen densidades de probabilidad de encontrar uno de los electrones en las coordenadas \vec{x}_1 y otro en las coordenadas \vec{x}_2 , independientemente de las coordenadas de los restantes: $\sigma(\vec{x}_1, \vec{x}_2)$. En efecto, la densidad de probabilidad de encontrar el electrón 1 en \vec{x}_1 y, simultáneamente, el electrón 2 en \vec{x}_2 es

$$\begin{aligned} \sigma_{12}(\vec{x}_1, \vec{x}_2) &= \int |\Psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots \vec{x}_N)|^2 d\vec{x}_3 \dots d\vec{x}_N \\ &= \rho_{12}(\vec{x}_1, \vec{x}_2; \vec{x}_1, \vec{x}_2) \end{aligned}$$

y la densidad de probabilidad de encontrar el electrón 2 en \vec{x}_1 y, simultáneamente, el electrón 1 en \vec{x}_2 será igual a aquélla, ya que dichas partículas son indistinguibles. El mismo razonamiento puede extenderse a las $N(N-1)/2$ parejas de electrones, de modo que

$$\begin{aligned} \sigma(\vec{x}_1, \vec{x}_2) &= [N(N-1)/2] 2 \sigma_{12}(\vec{x}_1, \vec{x}_2) \\ &= N(N-1) \rho_{12}(\vec{x}_1, \vec{x}_2; \vec{x}_1, \vec{x}_2) \\ &= \Gamma(\vec{x}_1, \vec{x}_2; \vec{x}_1, \vec{x}_2) . \end{aligned}$$

El valor esperado de un operador del tipo $\hat{G}(1, \dots N) = \sum_{i=1}^N \sum_{j>i}^N \hat{g}(i, j)$ con $\hat{g}(i, j)$ simétrico será

$$\begin{aligned} \langle G \rangle_{\Psi} &= [N(N-1)/2] \langle \Psi | \hat{g}(1, 2) \Psi \rangle \\ &= [N(N-1)/2] \text{Tr} \{ \hat{g}(1, 2) \hat{\rho}_{12}(1, 2) \} , \end{aligned}$$

o, introduciendo el operador $\hat{\Gamma}$ asociado a la matriz $\Gamma(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2')$,*

* Algunos autores normalizan la matriz de densidad de orden 2 a $N(N-1)/2$, de forma que no aparezca el factor $1/2$ en la expresión de $\langle G \rangle$.

$$\langle G \rangle_{\Psi} = \frac{1}{2} \text{Tr} (\hat{g} \hat{\Gamma}) .$$

Si desarrollamos la traza en la representación de posiciones

$$\langle G \rangle_{\Psi} = (1/2) \int \langle \vec{x}_1, \vec{x}_2 | \hat{g} \hat{\Gamma} | \vec{x}_1, \vec{x}_2 \rangle d\vec{x}_1 d\vec{x}_2$$

y suponemos que el operador \hat{g} es local en dicha representación:

$$\langle \vec{x}_1', \vec{x}_2' | \hat{g}(1, 2) \psi(1, 2) \rangle = \hat{g}(\vec{x}_1', \vec{x}_2') \langle \vec{x}_1', \vec{x}_2' | \psi(1, 2) \rangle ,$$

podremos poner

$$\langle G \rangle_{\Psi} = (1/2) \int (\hat{g}(\vec{x}_1', \vec{x}_2') \Gamma(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2))_{\vec{x}_1'=\vec{x}_1, \vec{x}_2'=\vec{x}_2} d\vec{x}_1 d\vec{x}_2 .$$

Para tratar operadores bielectrónicos que no dependan del spin: $\hat{g}(\vec{x}_1', \vec{x}_2') = \hat{g}(\vec{r}_1', \vec{r}_2')$, convendrá introducir una *matriz de densidad reducida de orden 2 sin spin*:

$$\begin{aligned} P(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) &= \text{Tr}_{\omega_1, \omega_2} \Gamma(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2) \\ &= \int \Gamma(\vec{r}_1' \omega_1, \vec{r}_2' \omega_2; \vec{r}_1 \omega_1, \vec{r}_2 \omega_2) d\omega_1 d\omega_2 , \end{aligned}$$

en función de la cual

$$\langle G \rangle_{\Psi} = (1/2) \int (\hat{g}(\vec{r}_1', \vec{r}_2') P(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2))_{\vec{r}_1'=\vec{r}_1, \vec{r}_2'=\vec{r}_2} d\vec{r}_1 d\vec{r}_2$$

y si, además, el operador es diagonal en la representación de posiciones,

$$\langle G \rangle_{\Psi} = (1/2) \int \hat{g}(\vec{r}_1, \vec{r}_2) \sigma(\vec{r}_1, \vec{r}_2) d\vec{r}_1 d\vec{r}_2 ,$$

donde $\sigma(\vec{r}_1, \vec{r}_2) \equiv P(\vec{r}_1, \vec{r}_2; \vec{r}_1, \vec{r}_2)$ es la densidad de probabilidad de encontrar un electrón en \vec{r}_1 y otro en \vec{r}_2 , independientemente de su spin y de las posiciones y spines de los restantes electrones.

Ejercicio 3.4. Comprueba que

$$\gamma(\vec{x}_1; \vec{x}_1') = (N-1)^{-1} \int \Gamma(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2) d\vec{x}_2$$

y que

$$P(\vec{r}_1; \vec{r}_1') = (N-1)^{-1} \int P(\vec{r}_1, \vec{r}_2; \vec{r}_1', \vec{r}_2) d\vec{r}_2 .$$

Como ejemplo de utilización de las matrices de densidad reducidas, expresaremos en función de ellas la energía electrónica no relativista de una molécula con N electrones y M núcleos. El correspondiente hamiltoniano consta de términos mono- y bi-electrónicos que no dependen de las variables de spin; en representación de posiciones y unidades atómicas su forma es:

$$\hat{H} = \sum_i^N (-1/2) \nabla_i^2 - \sum_A^M \sum_i^N Z_A r_{iA}^{-1} + \sum_{j>i}^N r_{ij}^{-1}$$

donde Z_A es el número atómico del núcleo A, y el valor esperado de este operador sobre cualquier función de onda N-electrónica será

$$E = (-1/2) \int \{\nabla'^2 P(\vec{r}'; \vec{r})\}_{\vec{r}'=\vec{r}} d\vec{r} - \sum_A^M Z_A \int r_{1A}^{-1} \sigma(\vec{r}_1) d\vec{r}_1 + (1/2) \int r_{12}^{-1} \sigma(\vec{r}_1, \vec{r}_2) d\vec{r}_1 d\vec{r}_2$$

Observemos que este resultado es independiente del tipo de función N-electrónica utilizado, la cual puede llegar a ser extremadamente compleja si se pretende describir con precisión el sistema. Cabe destacar también que, de los tres términos en que aparece desglosada la energía electrónica, los dos últimos coinciden, respectivamente, con la energía clásica de atracción electrostática entre la distribución continua de carga $\sigma(\vec{r})$ y los núcleos, y con la repulsión electrostática clásica promediada entre dos electrones cuya distribución espacial de probabilidades viene dada por la función de correlación de par $\sigma(\vec{r}_1, \vec{r}_2)$.

Es posible definir matrices de densidad reducidas de órdenes cualesquiera entre 1 y N pero, dado que los observables que nos interesan constan únicamente de términos mono- y bi-electrónicos, no necesitaremos matrices de orden superior a 2. Esto tiene una notable implicación: aunque la función de onda de un sistema con más de dos electrones contiene más información que las matrices de densidad reducidas de orden 1 y 2, el exceso de información nos es irrelevante, y sería de gran interés poder obtener aquellas matrices sin pasar por el cálculo previo de la función de onda. La economía de esfuerzo que esto podría suponer deriva del hecho de que la función de onda depende de $4N$ variables, en tanto que la matriz de densidad de orden 2 (a partir de la cual se obtiene fácilmente la de orden 1, tal como se indica en el ejercicio 3.4) depende de 16 variables, independientemente del número de electrones que posea el sistema, cifra que se reduce a 6 si no estamos interesados en observables que dependan del spin. Por desgracia, no se conocen condiciones necesarias y suficientes que determinen dicha matriz salvo en casos de limitado interés, como es el de una función de onda monodeterminantal, cuyas matrices de densidad reducidas contienen la misma información que la función de onda. En efecto, se puede demostrar que, en este caso, la matriz de primer orden determina las de orden superior (en el ejercicio 3.5 se obtendrá la de orden 2 a partir de la de primer orden), lo cual concuerda con el carácter de partículas independientes de la función utilizada. No obstante, se sigue trabajando en el desarrollo de métodos para la obtención directa de la matriz de densidad de orden 2, los cuales podrían representar un importante salto cualitativo en la aplicabilidad de la química cuántica al estudio de sistemas con muchos electrones.

3.3 Matrices de densidad reducidas de sistemas polielectrónicos en bases discretas

Aunque la química cuántica se desarrolla habitualmente en la representación de posiciones, la mayor parte de los programas de cálculo recurren a un conjunto finito de funciones mono-

electrónicas para expresar los spin-orbitales, lo que equivale a utilizar una base discreta truncada. La conexión entre las representaciones discreta y continua de los operadores de densidad reducidos se obtiene de forma inmediata con la ayuda de las oportunas resoluciones de la identidad. Así, la representación de la matriz de densidad de primer orden* en la base de spin-orbitales $\{\dots \psi_i \dots\}$ será

$$\begin{aligned} (\gamma_\psi)_{ij} &= \langle \psi_i | \hat{\gamma} | \psi_j \rangle \\ &= \int \langle \psi_i | \vec{x} \rangle \langle \vec{x} | \hat{\gamma} | \vec{x}' \rangle \langle \vec{x}' | \psi_j \rangle d\vec{x} d\vec{x}' \\ &= \int \psi_i^*(\vec{x}) \gamma(\vec{x}; \vec{x}') \psi_j(\vec{x}') d\vec{x} d\vec{x}' \end{aligned}$$

y, si el conjunto $\{\dots \psi_i \dots\}$ es ortonormal y completo, podemos escribir también:

$$\begin{aligned} \gamma(\vec{x}; \vec{x}') &= \langle \vec{x} | \hat{\gamma} | \vec{x}' \rangle \\ &= \sum_{ij} \langle \vec{x} | \psi_i \rangle \langle \psi_i | \hat{\gamma} | \psi_j \rangle \langle \psi_j | \vec{x}' \rangle \\ &= \boldsymbol{\psi}(\vec{x}) \boldsymbol{\gamma}_\psi \boldsymbol{\psi}^\dagger(\vec{x}') . \end{aligned}$$

Con la matriz de segundo orden podemos proceder de forma análoga:

$$\begin{aligned} (\Gamma_\psi)_{ij,kl} &= \langle \psi_i \psi_j | \hat{\Gamma} | \psi_k \psi_l \rangle \\ &= \int \psi_i^*(\vec{x}_1) \psi_j^*(\vec{x}_2) \Gamma(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2') \psi_k(\vec{x}_1') \psi_l(\vec{x}_2') d\vec{x}_1 d\vec{x}_2 d\vec{x}_1' d\vec{x}_2' \end{aligned}$$

y

$$\Gamma(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2') = \sum_{ijkl} \psi_i(\vec{x}_1) \psi_j(\vec{x}_2) (\Gamma_\psi)_{ij,kl} \psi_k^*(\vec{x}_1') \psi_l^*(\vec{x}_2')$$

Las expresiones, en función de estas matrices, de los valores esperados de observables de los tipos F y G definidos en la sección anterior será

$$\begin{aligned} \langle F \rangle &= \text{Tr}(\mathbf{f} \boldsymbol{\gamma}_\psi) \\ &= \sum_{ij} f_{ij} (\boldsymbol{\gamma}_\psi)_{ji} \\ \langle G \rangle &= (1/2) \text{Tr}(\mathbf{g} \boldsymbol{\Gamma}_\psi) \\ &= \sum_{ijkl} g_{ij,kl} (\boldsymbol{\Gamma}_\psi)_{kl,ij} \end{aligned}$$

donde

$$g_{ij,kl} = \langle \psi_i \psi_j | \hat{g} | \psi_k \psi_l \rangle .$$

3.3.1 Funciones de onda multideterminantales

Un tipo muy general de función de onda polieletrónica es

$$\Psi = \sum_I C_I \Psi_I ,$$

donde cada Ψ_I es un determinante de Slater construido a partir de un conjunto ortonormal de

* En adelante omitiremos el calificativo "reducidas" para referirnos a las representaciones de los operadores $\hat{\gamma}$, $\hat{\Gamma}$ y \hat{P} cuando ello no induzca a confusión.

spin-orbitales $\{\dots \psi_i \dots\}$ y supondremos que Ψ está normalizada. Los elementos de la correspondiente matriz de densidad de primer orden expresada en la base $\{\dots \psi_i \dots\}$ guardan una relación sencilla con los coeficientes C_I . En efecto,

$$\begin{aligned} (\gamma_\psi)_{ij} &= \int \psi_i^*(\vec{x}_1) \gamma(\vec{x}_1; \vec{x}_1') \psi_j(\vec{x}_1') d\vec{x}_1 d\vec{x}_1' \\ &= N \int \psi_i^*(\vec{x}_1) \Psi(\vec{x}_1, \vec{x}_2 \dots \vec{x}_N) \Psi^*(\vec{x}_1', \vec{x}_2 \dots \vec{x}_N) \psi_j(\vec{x}_1') d\vec{x}_1 d\vec{x}_1' d\vec{x}_2 \dots d\vec{x}_N \\ &= N \sum_I \sum_J C_I C_J^* (N!)^{-1} \times \\ &\int \psi_i^*(\vec{x}_1) \sum_\alpha (-1)^{\pi_\alpha} \widehat{\mathcal{P}}_\alpha \{\psi_{1_1}(\vec{x}_1) \dots \psi_{N_1}(\vec{x}_N)\} \sum_\beta (-1)^{\pi_\beta} \widehat{\mathcal{P}}_\beta \{\psi_{1_j}^*(\vec{x}_1') \dots \psi_{N_j}^*(\vec{x}_N)\} \psi_j(\vec{x}_1') d\vec{x}_1 d\vec{x}_1' \dots d\vec{x}_N \end{aligned}$$

Si hemos dispuesto cada pareja de determinantes $\{\Psi_I, \Psi_J\}$ en máxima coincidencia y de forma que, si el spin-orbital ψ_i aparece en Ψ_I o el ψ_j aparece en Ψ_J , ocupen la primera columna (o fila) del determinante, la integral anterior se anulará a no ser que $\psi_{1_1} = \psi_i, \psi_{1_j} = \psi_j, \psi_{2_1} = \psi_{2_j}, \dots, \psi_{N_1} = \psi_{N_j}, \widehat{\mathcal{P}}_\alpha = \widehat{\mathcal{P}}_\beta$ y esta permutación sea tal que los spin-orbitales ψ_{1_1} y ψ_{1_j} aparezcan en la primera posición (es decir, con las coordenadas \vec{x}_1 y \vec{x}_1' , respectivamente). Como consecuencia, los determinantes Ψ_I y Ψ_J han de diferir, como máximo, en su primer spin-orbital para que den una contribución no nula $(\gamma_\psi)_{ij}$, spin-orbital que habrá de ser ψ_i para el primer determinante y ψ_j para el segundo. Por otra parte, el número de permutaciones que se pueden formar manteniendo fija la posición de un spin-orbital es $(N-1)!$, de modo que

$$(\gamma_\psi)_{ij} = \sum_{I, \psi_i \in I} C_I C_{I_j}^*,$$

donde el sumatorio se extiende a todos los determinantes de Slater en los que aparece el spin-orbital ψ_i y C_{I_j} es el coeficiente del determinante que se obtiene sustituyendo ψ_i por ψ_j en Ψ_I . En particular, cada elemento diagonal representa la contribución total a Ψ de los determinantes que contienen un spin-orbital determinado, y recibe el nombre de *número de ocupación de dicho spin-orbital en la función de onda Ψ* :

$$(\gamma_\psi)_{ii} = \sum_{I, \psi_i \in I} |C_I|^2,$$

de donde se deduce que

$$0 \leq (\gamma_\psi)_{ii} \leq 1.$$

Introduciendo el *número de ocupación n_{iI} del spin-orbital ψ_i en el determinante Ψ_I* como un número que toma los valores 0 ó 1 según el spin-orbital aparezca o no en el determinante (caso particular de la definición anterior para una función de onda monodeterminantal), podemos poner

$$(\gamma_\psi)_{ii} = \sum_I n_{iI} |C_I|^2.$$

De acuerdo con el enunciado del ejercicio 1.27, el espacio de Hilbert de estados monoeléctricos es producto directo del espacio que correspondería a un electrón sin spin y el espacio bidimensional engendrado por los estados de spin α y β . Como consecuencia, podemos escoger los spin-orbitales como producto de un orbital de un conjunto ortonormal $\{\dots \phi_i(\vec{r}) \dots\}$ por una de las funciones de spin $\alpha(\omega)$ o $\beta(\omega)$, lo que permite obtener una expresión explícita

para los elementos de la matriz de densidad de primer orden sin spin sumando los oportunos elementos de γ_ψ respecto de la variable de spin:

$$\begin{aligned} (\mathbf{P}_\phi)_{ij} &= (\text{Tr}_\omega \boldsymbol{\gamma})_{ij} \\ &= \langle \phi_i \alpha | \hat{\gamma} | \phi_j \alpha \rangle + \langle \phi_i \beta | \hat{\gamma} | \phi_j \beta \rangle \\ &= \sum_{I, \phi_i \alpha \in I} C_I C_{I_i}^* + \sum_{I, \phi_i \beta \in I} C_I C_{I_i}^* \end{aligned}$$

donde los índices 'i' y 'j' representan ahora orbitales y C_{I_i} es el coeficiente del determinante que se obtiene sustituyendo $\phi_i \alpha$ por $\phi_j \alpha$ en Ψ_I , si se trata del primer sumatorio, y $\phi_i \beta$ por $\phi_j \beta$ en el caso del segundo. Los elementos diagonales representarán *números de ocupación de cada orbital en la función de onda total Ψ* , y pueden expresarse en función de los *números de ocupación n_{iI} del orbital ϕ_i en el determinante Ψ_I ($n_{iI} \equiv n_{\phi_i \alpha I} + n_{\phi_i \beta I}$):*

$$\begin{aligned} (\mathbf{P}_\phi)_{ii} &= \sum_I (n_{\phi_i \alpha I} + n_{\phi_i \beta I}) |C_I|^2 \\ &= \sum_I n_{iI} |C_I|^2 . \end{aligned}$$

Los índices n_{iI} pueden tomar los valores 0, 1 ó 2, de forma que

$$0 \leq (\mathbf{P}_\phi)_{ii} \leq 2 .$$

Como en el caso del operador $\hat{\gamma}$, podemos obtener la representación de posiciones del operador \hat{P} a partir de su representación en la base ortonormal $\{\dots \phi_i \dots\}$ introduciendo las oportunas resoluciones de la identidad en $P(\vec{r}; \vec{r}')$:

$$\begin{aligned} P(\vec{r}; \vec{r}') &= \langle \vec{r} | \hat{P} | \vec{r}' \rangle \\ &= \sum_{ij} \langle \vec{r} | \phi_i \rangle \langle \phi_i | \hat{P} | \phi_j \rangle \langle \phi_j | \vec{r}' \rangle \\ &= \sum_{ij} \phi_i(\vec{r}) (\mathbf{P}_\phi)_{ij} \phi_j^*(\vec{r}') . \end{aligned}$$

y viceversa.

3.3.2 Funciones de onda monodeterminantales

Para una función de onda monodeterminantal

$$\Psi = |\psi_1 \dots \psi_N| ,$$

podemos aplicar los resultados obtenidos para funciones multideterminantales tomando $C_I = 1$ para $\Psi_I = |\psi_1 \dots \psi_N|$ y $C_J = 0$ para todo $J \neq I$. La matriz de densidad de primer orden en una base ortonormal de spin-orbitales cuyos N primeros elementos sean $\{\psi_1, \dots \psi_N\}$ tendrá todos sus elementos nulos excepto los N primeros elementos diagonales que tomarán el valor 1. Su expresión en representación de posiciones puede obtenerse utilizando la ley de transformación deducida al comienzo de esta sección:

$$\gamma(\vec{x}; \vec{x}') = \sum_a^N \psi_a(\vec{x}) \psi_a^*(\vec{x}') .$$

La densidad de probabilidad $\sigma(\vec{x})$ se reduce en este caso a la suma de las contribuciones de los spin-orbitales ocupados:

$$\sigma(\vec{x}) = \sum_a^N |\psi_a(\vec{x})|^2 .$$

Ejercicio 3.5. Expresa la matriz de densidad reducida de orden 2 asociada a la función de onda N-electrónica monodeterminantal $|\psi_1 \dots \psi_N|$ en función de los spin-orbitales $\{\psi_1 \dots \psi_N\}$ y demuestra que, para dicha función de onda, se puede obtener aquella matriz de densidad a partir de la de orden 1 mediante la siguiente relación:

$$\Gamma(1, 2; 1', 2') = \gamma(1;1') \gamma(2;2') - \gamma(2;1') \gamma(1;2') ,$$

donde $1 \equiv \vec{x}_1$ y $2 \equiv \vec{x}_2$.

Si los spin-orbitales del determinante son productos de orbitales ortonormales, $\phi_a(\vec{r})$, por una de las funciones de spin $\alpha(\omega)$ o $\beta(\omega)$, podemos completar formalmente el conjunto de orbitales $\{\phi_a\}$ con cualquier base ortonormal de su complemento ortogonal y representar el operador \hat{P} en la base resultante:

$$(\mathbf{P}_\phi)_{ij} = n_i \delta_{ij} ,$$

donde n_i es el número de ocupación del orbital ϕ_i en el determinante. Pasando a la representación de posiciones:

$$\begin{aligned} P(\vec{r}; \vec{r}') &= \sum_{ij} \phi_i(\vec{r}) n_i \delta_{ij} \phi_j(\vec{r}') \\ &= \sum_i \phi_i(\vec{r}) n_i \phi_i(\vec{r}') \\ &= \sum_{a,\alpha}^{N^\alpha} \phi_a(\vec{r}) \phi_a^*(\vec{r}') + \sum_{a,\beta}^{N^\beta} \phi_a(\vec{r}) \phi_a^*(\vec{r}') \end{aligned}$$

donde N^α y N^β son los números de spin-orbitales con función de spin α y β , respectivamente que contiene el determinante.

En un determinante *restringido de capas cerradas*, los números de ocupación tomarán los valores 2 ó 0 según se refieran a un orbital ocupado o virtual, respectivamente, y podemos escribir

$$P(\vec{r}; \vec{r}') = \sum_a^{N/2} 2 \phi_a(\vec{r}) \phi_a^*(\vec{r}') .$$

En un determinante *no restringido* los spin-orbitales α y los β tienen, en general distintas partes espaciales (que designaremos como ϕ_a^α y ϕ_a^β , respectivamente) no necesariamente ortogonales entre sí:

$$\langle \phi_a^\alpha | \phi_b^\beta \rangle \neq \delta_{ab}$$

si bien los conjuntos $\{\dots \phi_a^\alpha \dots\}$ y $\{\dots \phi_a^\beta \dots\}$ son ortonormales:

$$\langle \phi_a^\alpha | \phi_b^\alpha \rangle = \langle \phi_a^\beta | \phi_b^\beta \rangle = \delta_{ab} ,$$

de forma que los spin-orbitales del determinante siguen formando un conjunto ortonormal.

La traza parcial respecto de ω de la matriz $\gamma(\vec{x}; \vec{x}')$ conduce ahora a la siguiente expresión para $P(\vec{r}; \vec{r}')$:

$$\begin{aligned}
 P(\vec{r}; \vec{r}') &= \text{Tr}_\omega \gamma(\vec{x}; \vec{x}') \\
 &= \gamma(\vec{r}\alpha; \vec{r}'\alpha) + \gamma(\vec{r}\beta; \vec{r}'\beta) \\
 &= \sum_a^N \psi_a(\vec{r}\alpha) \psi_a^*(\vec{r}'\alpha) + \sum_a^N \psi_a(\vec{r}\beta) \psi_a^*(\vec{r}'\beta) \\
 &= \sum_a^N \langle \vec{r}\alpha | \phi_a^{\alpha} \rangle \langle \phi_a^{\alpha} | \vec{r}'\alpha \rangle + \sum_a^N \langle \vec{r}\beta | \phi_a^{\beta} \rangle \langle \phi_a^{\beta} | \vec{r}'\beta \rangle \\
 &= \sum_a^{N^\alpha} \phi_a^\alpha(\vec{r}) (\phi_a^\alpha)^*(\vec{r}') + \sum_a^{N^\beta} \phi_a^\beta(\vec{r}) (\phi_a^\beta)^*(\vec{r}') .
 \end{aligned}$$

Definiendo las matrices de spin α y de spin β como

$$\begin{aligned}
 P^{\alpha,\alpha}(\vec{r}; \vec{r}') &\equiv \gamma(\vec{r} \alpha; \vec{r}' \alpha) \\
 &= \sum_a^{N^\alpha} \phi_a^\alpha(\vec{r}) (\phi_a^\alpha)^*(\vec{r}')
 \end{aligned}$$

y

$$\begin{aligned}
 P^{\beta,\beta}(\vec{r}; \vec{r}') &\equiv \gamma(\vec{r} \beta; \vec{r}' \beta) \\
 &= \sum_a^{N^\beta} \phi_a^\beta(\vec{r}) (\phi_a^\beta)^*(\vec{r}') ,
 \end{aligned}$$

respectivamente, podemos expresar $P(\vec{r}; \vec{r}')$ en la forma

$$P(\vec{r}; \vec{r}') = P^{\alpha,\alpha}(\vec{r}; \vec{r}') + P^{\beta,\beta}(\vec{r}; \vec{r}') .$$

La diferencia entre estas dos matrices se conoce como *matriz de densidad de spin*:

$$P^S(\vec{r}; \vec{r}') \equiv P^{\alpha,\alpha}(\vec{r}; \vec{r}') - P^{\beta,\beta}(\vec{r}; \vec{r}')$$

y su diagonal

$$\sigma^S(\vec{r}) \equiv P^S(\vec{r}; \vec{r})$$

nos da idea de la asimetría en la contribución de los electrones α y β a la densidad electrónica total.

La matriz que representa al operador \hat{P} en una base de orbitales que contenga a los ocupados se puede seguir calculando como:

$$\begin{aligned}
 (\mathbf{P}_\phi)_{ij} &= \langle \phi_i^{\alpha} | \hat{P} | \phi_j^{\alpha} \rangle \\
 &= (\text{Tr}_\omega \boldsymbol{\gamma})_{ij} \\
 &= \langle \phi_i^{\alpha} | \hat{\gamma} | \phi_j^{\alpha} \rangle + \langle \phi_i^{\beta} | \hat{\gamma} | \phi_j^{\beta} \rangle \\
 &= \delta_{ij} n(\phi_i^{\alpha}) .
 \end{aligned}$$

Ejercicio 3.6. Comprueba que, para una función de onda monodeterminantal cuyos spin-orbitales son productos de un orbital por una de las funciones de spin α o β ,

$$\gamma(\vec{x}; \vec{x}') = \alpha(\omega) \alpha^*(\omega') P^{\alpha,\alpha}(\vec{r}; \vec{r}') + \beta(\omega) \beta^*(\omega') P^{\beta,\beta}(\vec{r}; \vec{r}')$$

3.4 Spin-orbitales naturales

La hermiticidad de la matriz de densidad de primer orden $\boldsymbol{\gamma}_\psi$ de un sistema polielectrónico en una base ortonormal de vectores de estado monoeléctronicos $\{\dots \psi_i \dots\}$ hace que dicha matriz sea diagonalizable mediante una transformación unitaria $\boldsymbol{\eta} = \boldsymbol{\psi} \mathbf{U}$, es decir:

$$\mathbf{U}^\dagger \boldsymbol{\gamma}_\psi \mathbf{U} = \boldsymbol{\gamma}_\eta,$$

con $(\boldsymbol{\gamma}_\eta)_{ij} = n_i \delta_{ij}$ y $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. Los spin-orbitales asociados a la nueva base se conocen como *spin-orbitales naturales* y el desarrollo de $\gamma(\vec{x}; \vec{x}')$ en función de éstos se reduce a

$$\begin{aligned} \gamma(\vec{x}; \vec{x}') &= \boldsymbol{\eta}(\vec{x}) \boldsymbol{\gamma}_\eta \boldsymbol{\eta}^\dagger(\vec{x}') \\ &= \sum_i \eta_i(\vec{x}) n_i \eta_i^*(\vec{x}'). \end{aligned}$$

La descomposición espectral del operador $\hat{\gamma}^*$ en la nueva base

$$\hat{\gamma} = \sum_i n_i |\eta_i\rangle \langle \eta_i|$$

revela una interesante interpretación física de los spin-orbitales naturales y de sus números de ocupación. Dividiendo por N dicha descomposición se obtiene un operador de densidad reducido monoeléctrico que representa un estado mezcla estadística de los estados monoeléctricos η_i con probabilidades n_i/N :

$$\hat{\rho}_1 = \hat{\gamma}/N = \sum_i (n_i/N) |\eta_i\rangle \langle \eta_i|.$$

Por lo tanto, aunque, en general, no podamos asociar un vector de estado a un electrón de un sistema polielectrico, es posible asociarle un operador densidad que representa un estado mezcla de spin-orbitales naturales, siendo la probabilidad de que el electrón ocupe uno de dichos estados puros monoeléctricos proporcional al correspondiente número de ocupación.

En el caso particular de que la función de onda que describe el estado puro sea un determinante de Slater, $\hat{\gamma}$ es el operador de proyección* sobre el subespacio engendrado por los spin-orbitales ocupados

$$\hat{\gamma} = \sum_a^N |\eta_a\rangle \langle \eta_a|,$$

y se demuestra que el recíproco también se cumple. El estado que representa $\hat{\rho}_1$ es, en este caso, una mezcla equiprobable de los N estados monoeléctricos η_1, \dots, η_N . La posibilidad de expresar la densidad de probabilidad electrónica como suma de contribuciones de cada orbital ocupado (apartado 3.3.2) puede considerarse como una consecuencia directa de lo anterior, y puede extenderse a funciones polielectricas cualesquiera utilizando los spin-orbitales naturales:

$$\begin{aligned} \sigma(\vec{x}) &= \boldsymbol{\eta}(\vec{x}) \boldsymbol{\gamma}_\eta \boldsymbol{\eta}^\dagger(\vec{x}) \\ &= \sum_i n_i |\psi_i(\vec{x})|^2. \end{aligned}$$

Los spin-orbitales naturales tienen propiedades interesantes de cara a la construcción de funciones de onda polielectricas multideterminantales. Para analizar este punto, empezaremos por demostrar un teorema según el cual la suma de r elementos diagonales cualesquiera de una matriz hermítica es siempre menor o igual que la suma de de sus r mayores valores propios. En

* Se demuestra que su espectro es discreto.

* Aunque $\hat{\gamma}$ sea un operador de proyección no representa un estado puro, ya que no proyecta sobre un espacio monodimensional; de hecho, ni siquiera es un operador densidad en el sentido definido en el capítulo 1, ya su traza no es igual a la unidad (para $N > 1$).

efecto, utilizando la relación

$$\Upsilon_{\psi} = \mathbf{U} \Upsilon_{\eta} \mathbf{U}^{\dagger}$$

se obtiene

$$\begin{aligned} \sum_{i=1}^r \Upsilon_{ii} &= \sum_{i=1}^r \sum_{j,k=1}^n U_{ij} n_j \delta_{jk} U_{ik}^* \\ &= \sum_{j=1}^r n_j \sum_{i=1}^r |U_{ij}|^2 \\ &= \sum_{j=1}^r n_j \sum_{i=1}^r |U_{ij}|^2 + \sum_{j=r+1}^n n_j \sum_{i=1}^r |U_{ij}|^2 \\ &= \sum_{j=1}^r n_j (1 - \sum_{i=r+1}^n |U_{ij}|^2) + \sum_{j=r+1}^n n_j \sum_{i=1}^r |U_{ij}|^2, \end{aligned}$$

ya que $\sum_{i=1}^n |U_{ij}|^2 = 1$. Adoptando para los valores propios la ordenación de mayor a menor:

$n_1 \geq \dots \geq n_r \dots \geq n_n$ podemos poner

$$\sum_{i=1}^r \Upsilon_{ii} \leq \sum_{j=1}^r n_j - n_r \sum_{j=1}^r \sum_{i=r+1}^n |U_{ij}|^2 + n_{r+1} \sum_{i=1}^r \sum_{j=r+1}^n |U_{ij}|^2,$$

donde el signo igual solo rige en el improbable caso de que los vectores propios $\{\eta_1, \dots, \eta_r\}$ sean degenerados entre si y ocurra otro tanto con $\{\eta_{r+1}, \dots, \eta_n\}$, o cuando la transformación \mathbf{U}

afecta únicamente a los spin-orbitales ψ_r y ψ_{r+1} . Utilizando de nuevo la unitariedad de \mathbf{U} ,

$$\begin{aligned} \sum_{j=1}^r \sum_{i=r+1}^n |U_{ij}|^2 &= \sum_{j=1}^r (1 - \sum_{i=1}^r |U_{ij}|^2) \\ &= r - \sum_{j=1}^r \sum_{i=1}^r |U_{ij}|^2 \end{aligned}$$

y

$$\begin{aligned} \sum_{i=1}^r \sum_{j=r+1}^n |U_{ij}|^2 &= \sum_{i=1}^r (1 - \sum_{j=1}^r |U_{ij}|^2) \\ &= r - \sum_{i=1}^r \sum_{j=1}^r |U_{ij}|^2, \end{aligned}$$

de modo que

$$\sum_{i=1}^r \Upsilon_{ii} \leq \sum_{j=1}^r n_j + (n_{r+1} - n_r) (r - \sum_{i=1}^r \sum_{j=1}^r |U_{ij}|^2)$$

finalmente, las relaciones

$$\sum_{i=1}^r \sum_{j=1}^r |U_{ij}|^2 \leq \sum_{i=1}^r \sum_{j=1}^n |U_{ij}|^2 = r$$

y

$$n_{r+1} \leq n_r$$

conducen a*

$$\sum_{i=1}^r \Upsilon_{ii} \leq \sum_{j=1}^r n_j.$$

Una función de onda N-electrónica multideterminantal se puede expresar como combinación lineal de determinantes de Slater contruidos a partir de spin-orbitales de cualquiera de las dos bases:

$$\begin{aligned} \Psi &= \sum_I C_I |\psi_{1_I} \dots \psi_{N_I}| \\ &= \sum_I D_I |\eta_{1_I} \dots \eta_{N_I}| \end{aligned}$$

y, teniendo en cuenta que

$$\sum_{I/\psi_i \in I} |C_I|^2 = \Upsilon_{ii}$$

* En la aplicación que haremos de este resultado nunca se dará el caso $n_{r+1} = n_r$, ya que en una función multideterminantal siempre se incluyen conjuntos completos de spin-orbitales degenerados.

$$y \sum_{I/\eta_i \in I} |D_I|^2 = n_i,$$

podemos utilizar la desigualdad anteriormente deducida para establecer la siguiente relación:

$$\sum_{i=1}^r \sum_{I/\eta_i \in I} |D_I|^2 \geq \sum_{i=1}^r \sum_{I/\psi_i \in I} |C_I|^2,$$

es decir, los determinantes de Slater en los que figuran los r spin-orbitales naturales de mayores números de ocupación tienen un peso en Ψ mayor o igual que los determinantes que contienen cualquier conjunto de r spin-orbitales de cualquier otra base monoeléctronica. De acuerdo con esto, una base truncada de spin-orbitales naturales permitirá construir funciones multideterminantales de mayor calidad que las obtenidas a partir de otros conjuntos de spin-orbitales con igual número de elementos.

Esta interesante propiedad de los spin-orbitales naturales nos lleva, de nuevo, al problema del cálculo directo de matrices de densidad reducidas, ya que sirve de poco disponer de unos orbitales óptimos para la construcción de desarrollos multideterminantales reducidos si debemos calcular una función polieletrónica más compleja como paso previo para la obtención de aquellos orbitales. No obstante, se han propuesto métodos de cálculo de spin-orbitales pseudo-naturales que no requieren dicho paso previo. Por ejemplo, puede obtenerse una primera aproximación a los orbitales pseudo-naturales a partir de un desarrollo multideterminantal reducido y, seleccionando de entre aquellos orbitales los de mayores números de ocupación, construir un nuevo desarrollo multideterminantal de la misma dimensión que el primero. Los coeficientes de uno y otro desarrollo se optimizan variacionalmente, tal como se indica en la sección 2.2 (método de *interacción de configuraciones*). Este proceso puede repetirse de forma iterativa, si bien diverge con frecuencia después de algunas iteraciones. Un procedimiento alternativo que acostumbra dar buenos resultados consiste en obtener unos orbitales pseudo-naturales a partir de un desarrollo multideterminantal moderado y seleccionar los de mayores números de ocupación para construir una función con un número de determinantes superior.

Ejercicio 3.7. ¿Cuáles son los spin-orbitales naturales de una función de onda monodeterminantal?

Los *orbitales naturales* se definen como aquéllos que diagonalizan la matriz de densidad de primer orden sin spin y, si los vectores de estado monoeléctronicos de base son de la forma $\phi_i(\vec{r}) \alpha(\omega)$ ó $\phi_i(\vec{r}) \beta(\omega)$, podemos obtener aquellos orbitales integrando respecto de la variable de spin los spin-orbitales naturales, lo que permite extender a los primeros la discusión realizada para los segundos.

Los vectores que diagonalizan la matriz de densidad de orden 2 son estados bielectrónicos que, en la representación de posiciones, reciben el nombre de *geminales naturales*. Por el

momento, no se han encontrado para estas funciones aplicaciones del interés que presentan las de los spin-orbitales naturales.

3.5 Análisis poblacional

En el estudio de la estructura electrónica molecular, los orbitales ortonormales $\{\phi_1, \dots, \phi_i \dots\}$ que se utilizan para construir los determinantes de Slater son, normalmente, combinaciones lineales de unas funciones prefijadas de antemano $\{\chi_1, \dots, \chi_\mu, \dots\}$, las cuales suelen ser funciones gaussianas o de Slater, no necesariamente ortonormales, centradas en las posiciones de los núcleos atómicos:

$$\boldsymbol{\phi} = \boldsymbol{\chi} \mathbf{C} .$$

La definición habitual de la matriz de densidad de primer orden en la base atómica es (vease, por ejemplo, McWeeny [89] p. 134, Christoffersen [89] p. 507 ó Szabo[89] p. 139):

$$\tilde{\mathbf{P}}_{\chi} = \mathbf{C} \mathbf{P}_{\phi} \mathbf{C}^{\dagger} ,$$

donde hemos utilizado una tilde para distinguir la matriz así definida de la representación según el producto escalar del operador $\hat{\mathbf{P}}$ en aquella base: \mathbf{P}_{χ} . Dicha definición se suele justificar mediante el siguiente "algoritmo" de cambio de base:

$$\begin{aligned} P(\vec{r}; \vec{r}') &= \boldsymbol{\phi}(\vec{r}) \mathbf{P}_{\phi} \boldsymbol{\phi}^{\dagger}(\vec{r}') \\ &= \boldsymbol{\chi}(\vec{r}) \mathbf{C} \mathbf{P}_{\phi} \mathbf{C}^{\dagger} \boldsymbol{\chi}^{\dagger}(\vec{r}') \\ &= \boldsymbol{\chi}(\vec{r}) \tilde{\mathbf{P}}_{\chi} \boldsymbol{\chi}^{\dagger}(\vec{r}') . \end{aligned}$$

$\tilde{\mathbf{P}}_{\chi}$ no corresponde a ninguno de los conceptos de representación matricial de operadores que hemos introducido en el capítulo 2. En efecto, teniendo en cuenta la forma que adopta la resolución de la identidad en una base discreta no ortonormal (apartado d) del ejercicio 2.1), podemos escribir

$$\begin{aligned} \langle \vec{r} | \hat{\mathbf{P}} | \vec{r}' \rangle &= \sum_{\mu\nu\lambda\sigma} \langle \vec{r} | \chi_{\mu} \rangle (\mathbf{S}_{\chi}^{-1})_{\mu\nu} \langle \chi_{\nu} | \hat{\mathbf{P}} \chi_{\lambda} \rangle (\mathbf{S}_{\chi}^{-1})_{\lambda\sigma} \langle \chi_{\sigma} | \vec{r}' \rangle \\ &= \boldsymbol{\chi}(\vec{r}) \mathbf{S}_{\chi}^{-1} \mathbf{P}_{\chi} \mathbf{S}_{\chi}^{-1} \boldsymbol{\chi}^{\dagger}(\vec{r}') \end{aligned}$$

donde \mathbf{P}_{χ} es la representación según el producto escalar de $\hat{\mathbf{P}}$ en la base $\boldsymbol{\chi}$. De las relaciones anteriores se desprende que

$$\tilde{\mathbf{P}}_{\chi} = \mathbf{S}_{\chi}^{-1} \mathbf{P}_{\chi} \mathbf{S}_{\chi}^{-1}$$

o, despejando \mathbf{P}_{χ} y utilizando la expresión de $\tilde{\mathbf{P}}_{\chi}$ en función de \mathbf{P}_{ϕ} ,

$$\begin{aligned} \mathbf{P}_{\chi} &= \mathbf{S}_{\chi} \tilde{\mathbf{P}}_{\chi} \mathbf{S}_{\chi} \\ &= \mathbf{S}_{\chi} \mathbf{C} \mathbf{P}_{\phi} \mathbf{C}^{\dagger} \mathbf{S}_{\chi} , \end{aligned}$$

expresión que podría haberse obtenido de forma más directa introduciendo los oportunos desarrollos de la identidad en $\langle \chi_{\mu} | \hat{\mathbf{P}} \chi_{\nu} \rangle$. Multiplicando la expresión anterior por la izquierda por

\mathbf{C}^\dagger y por la derecha por \mathbf{C} y teniendo en cuenta que la base ϕ es ortonormal (de modo que $\mathbf{S}_\phi = \mathbf{C}^\dagger \mathbf{S}_\chi \mathbf{C} = \mathbf{1}$) se recupera la ley de transformación correcta (ejercicio 2.2):

$$\mathbf{P}_\phi = \mathbf{C}^\dagger \mathbf{P}_\chi \mathbf{C} .$$

De acuerdo con las relaciones deducidas en el ejercicio 2.3, la traza de un operador es igual a la suma de los elementos diagonales de su representación matricial *cursiva* en cualquier base discreta, de modo que,

$$\begin{aligned} N &= \text{Tr } \hat{\mathbf{P}} \\ &= \text{Tr } \mathbf{P}_\chi \\ &= \text{Tr } \mathbf{S}_\chi^{-1} \mathbf{P}_\chi \\ &= \text{Tr } \tilde{\mathbf{P}}_\chi \mathbf{S}_\chi . \end{aligned}$$

De aquí que deba multiplicarse $\tilde{\mathbf{P}}_\chi$ por \mathbf{S}_χ para obtener una matriz cuyos elementos diagonales proporcionen una partición del número total de electrones entre las funciones de la base χ :

$$\begin{aligned} N &= \sum_\mu (\tilde{\mathbf{P}}_\chi \mathbf{S}_\chi)_{\mu\mu} \\ &= \sum_\mu p_\mu , \end{aligned}$$

donde hemos introducido la *población electrónica total* de Mulliken (Mulliken [55] y [62]) sobre el orbital de base χ_μ , p_μ , definida de la siguiente forma:

$$\begin{aligned} p_\mu &= (\tilde{\mathbf{P}}_\chi \mathbf{S}_\chi)_{\mu\mu} \\ &= (\mathbf{S}_\chi^{-1} \mathbf{P}_\chi)_{\mu\mu} \\ &= (\mathbf{C} \mathbf{P}_\phi \mathbf{C}^\dagger \mathbf{S}_\chi)_{\mu\mu} . \end{aligned}$$

La suma de éstas poblaciones extendida a las funciones de base centradas en un núcleo A es la población electrónica total sobre dicho átomo:

$$P_A = \sum_{\mu \in A} p_\mu$$

y, la suma de estas poblaciones atómicas extendida a todos los núcleos de la molécula es, como era de esperar, el número total de electrones:

$$N = \sum_A P_A .$$

Para funciones monodeterminantales $(\mathbf{P}_\phi)_{ij} = n_i \delta_{ij}$ (apartado 3.3.2) y la matriz $\tilde{\mathbf{P}}_\chi$ adopta una forma particularmente sencilla:

$$\begin{aligned} (\tilde{\mathbf{P}}_\chi)_{\mu\nu} &= \sum_{ij} c_{\mu i} n_i \delta_{ij} c_{\nu j}^* \\ &= \sum_a c_{\mu a} n_a c_{\nu a}^* . \end{aligned}$$

La población electrónica total sobre χ_μ será, en este caso,

$$p_\mu = \sum_a \sum_\nu c_{\mu a} n_a c_{\nu a}^* (\mathbf{S}_\chi)_{\nu\mu}$$

y, a partir de ella, se pueden definir diversas "poblaciones electrónicas" que permiten efectuar distintos desgloses de las poblaciones de Mulliken (vease, por ejemplo, Pilar [68], p. 499-500).

La expresión anterior para p_μ puede extenderse a funciones multideterminantales utilizando

orbitales naturales η_i , ya que, en este caso, también se cumple $(\mathbf{P}_\eta)_{ij} = n_i \delta_{ij}$ (sección 3.4) y

$$\begin{aligned} p_\mu &= (\mathbf{C} \mathbf{P}_\eta \mathbf{C}^\dagger \mathbf{S}_\chi)_{\mu\mu} \\ &= \sum_i \sum_v c_{\mu i} n_i c_{v i}^* (\mathbf{S}_\chi)_{v\mu} . \end{aligned}$$

La invarianza de la traza de un producto de matrices respecto de permutaciones cíclicas de éstas permite establecer nuevas particiones del número total de electrones:

$$\begin{aligned} N &= \text{Tr} (\mathbf{S}_\chi^\alpha \tilde{\mathbf{P}}_\chi \mathbf{S}_\chi^{1-\alpha}) \\ &= \text{Tr} (\mathbf{S}_\chi^{-1+\alpha} \mathbf{P}_\chi \mathbf{S}_\chi^{-\alpha}) . \end{aligned}$$

La más extendida, propuesta por *Löwdin* (Löwdin []), es la que se obtiene tomando $\alpha=1/2$:

$$\begin{aligned} N &= \text{Tr} (\mathbf{S}_\chi^{1/2} \tilde{\mathbf{P}}_\chi \mathbf{S}_\chi^{1/2}) \\ &= \text{Tr} (\mathbf{S}_\chi^{-1/2} \mathbf{P}_\chi \mathbf{S}_\chi^{-1/2}) \\ &= \text{Tr} (\mathbf{P}_{\chi'}) \end{aligned}$$

donde $\mathbf{P}_{\chi'}$ es la expresión matricial de $\hat{\mathbf{P}}$ en la base que se obtiene al aplicar el método de ortonormalización de Löwdin a χ (ejercicios 2.2 y 2.5):

$$\chi' = \chi \mathbf{S}_\chi^{-1/2} .$$

Es interesante señalar que, aunque $\tilde{\mathbf{P}}_\chi$ no sea la auténtica representación del operador $\hat{\mathbf{P}}$, se cumple que, para cualquier operador monoeléctrico que no dependa del spin, $\hat{f}(\vec{r})$,

$$\begin{aligned} \text{Tr} (\hat{\mathbf{P}} \hat{f}) &= \text{Tr} (\mathbf{P}_\chi \mathbf{f}_\chi) \\ &= \text{Tr} (\mathbf{S}_\chi^{-1} \mathbf{P}_\chi \mathbf{S}_\chi^{-1} \mathbf{f}_\chi) \\ &= \text{Tr} (\tilde{\mathbf{P}}_\chi \mathbf{f}_\chi) . \end{aligned}$$

Esto permite calcular el valor esperado de un operador del tipo $\hat{F}(\vec{r}_1 \dots \vec{r}_N) = \sum_i^N \hat{f}(\vec{r}_i)$ utilizando las matrices $\tilde{\mathbf{P}}_\chi$ y \mathbf{f}_χ como si se tratara de representaciones cursivas de los correspondientes operadores:

$$\langle \Psi | \hat{F} | \Psi \rangle = \text{Tr} (\tilde{\mathbf{P}}_\chi \mathbf{f}_\chi) .$$

No obstante, consideramos que este resultado es una coincidencia que puede inducir a errores, y creemos conveniente utilizar siempre las representaciones matriciales introducidas en el capítulo 2.