

ARTÍCULO

IMPACTO DEL SUPERCÓMPUTO SOBRE LA QUÍMICA CUÁNTICA

Jorge Garza Olguín
Departamento de Química,
Universidad Autónoma Metropolitana-Iztapalapa.

Resumen

El desarrollo vertiginoso del supercómputo ha impactado de manera importante varias áreas del conocimiento. En particular, la química cuántica se ha visto beneficiada notablemente por este crecimiento, permitiéndole estudiar sistemas electrónicos que hasta hace algunos años eran imposibles de ser estudiados por técnicas sofisticadas. En este trabajo revisamos brevemente el método de Hartree-Fock, el cual sirve como punto de partida de muchos métodos basados en la función de onda, además su implementación en códigos computacionales sirve también como punto de partida para otros métodos alternativos a la función de onda. Mencionaremos los puntos donde se ha visto beneficiada la química cuántica computacional por el hardware que existe hoy en día y por las técnicas de programación que se han adoptado para explotar al máximo este hardware.

Palabras clave: Supercómputo, química cuántica, Hartree-Fock, códigos computacionales.

Fundamentos de la química cuántica

En 1926 Erwin Shrödinger publicó una serie de trabajos, en alemán¹⁻⁶ y luego en inglés⁷, donde plantea una nueva forma de describir a la materia. En particular su énfasis se centra en los átomos, protones y electrones. Su propuesta se basó en las ideas revolucionarias de Louis de Broglie (cuyo nombre es abreviado ya que su nombre completo es Louis Victor Pierre Raymond duc de Broglie). Shrödinger cita textualmente⁷ "La teoría que está reportada en las siguientes páginas está basada sobre la muy interesante investigación fundamental de L. de Broglie sobre la que él llamó ondas de fase "*ondes de phase*" las cuales se piensa están asociadas con el movimiento de puntos materiales, especialmente con el movimiento de un electrón o un protón". Sin duda, estas nuevas ideas generaron gran revuelo en la comunidad científica e incluso se generaron corrientes filosóficas sobre el uso e interpretación de esta mecánica ondulatoria, en contraparte a la ya conocida mecánica clásica de Newton, Hamilton o Lagrange.⁸ Ya que con esta nueva mecánica es imposible determinar exactamente la posición y la velocidad de un electrón, a cambio se debe de hablar sobre las regiones más probables donde se puede encontrar a esta partícula.⁹ Dentro de los sistemas que reporta Shrödinger como aplicación de su nueva forma de describir a la materia se encuentra el átomo de hidrógeno, el cual consiste de un protón y un electrón. Este era un sistema importante de estudiar ya que experimentalmente se tenía información espectroscópica que ninguna teoría podía dar cuenta de sus resultados experimentales de una manera lógica y consistente. Este fue un cimienta importante a la mecánica ondulatoria ya que pudo reproducir muchos de los datos experimentales que se tenían a la mano y de paso dejaba claro el camino a seguir para describir a los sistemas electrónicos.

La ecuación de Shrödinger con la que se trabaja generalmente en la química cuántica es aquella donde no se tiene una dependencia explícita en el tiempo, su forma es

$$\hat{H}\Psi = E\Psi \quad (1)$$

\hat{H} representa a un operador, el operador de Hamilton, y contiene todas las interacciones posibles entre los núcleos y electrones en una molécula.⁹ Evidentemente \hat{H} será único para un átomo, una molécula o un sólido. Al saber el tipo de interacciones que existen dentro de un sistema \hat{H} queda determinado exactamente y lo que hay que hacer es encontrar a la función Ψ , la función de onda, que al aplicarle el operador \hat{H} da como resultado a la misma función multiplicada por la constante E , esta constante representa a la energía del sistema. La ecuación 1 representa un problema de valores propios.

Cuando pensamos en una molécula, por ejemplo el agua, debemos de considerar el número de variables de las que depende la función Ψ , las cuales se obtienen al considerar las coordenadas espaciales de cada partícula, tres núcleos (dos hidrógenos y un oxígeno) y diez electrones (dos por parte de los hidrógenos y ocho por parte del oxígeno). Como cada partícula tiene tres coordenadas (x,y,z) entonces el número de variables asociadas con la función de onda del agua será de 39. Además es necesario tomar en cuenta la variable de espín de cada partícula, esto es muy importante ya que las propiedades magnéticas de la materia se deben al espín que tiene asociado cada partícula. Así que además de las coordenadas (x,y,z) se debe de tomar en cuenta una variable más, w. Sin duda alguna resolver un problema con tantas variables es muy complejo. De hecho solamente se ha resuelto, la ecuación 1, de manera analítica para sistemas que contienen un electrón. Aunque estos sistemas no están involucrados directamente con la descripción de muchas reacciones químicas han permitido generar ideas para el tratamiento de sistemas con muchos electrones. Por ejemplo, en la solución del átomo de hidrógeno se acuñó el nombre de orbital (para no confundir su función

con la órbita que se puede dar en un sistema descrito por la mecánica clásica) para nombrar a la función que describe al electrón de este sistema.

Debido a que no es posible resolver de manera analítica la ecuación 1 es necesario recurrir a aproximaciones. Una de ellas es la aproximación de Born-Oppenheimer,¹⁰ donde los núcleos se consideran fijos dentro de una molécula. El razonamiento matemático de esta aproximación se basa en el hecho de que el movimiento de los núcleos es mucho menor a aquel observado por los electrones. De esta manera se considera solamente como variables a las coordenadas de los electrones. En el caso de la molécula del agua la función de onda dependerá de cuarenta variables (30 espaciales y 10 de espín). Es claro que aún con la aproximación de núcleos fijos el problema sigue siendo enorme y más si pensamos en sistemas de tamaño considerable como el fullereno (C₆₀ 60 átomos de carbón) o parte de una proteína.

El método de Hartree-Fock

Dentro de la aproximación de núcleos fijos existen varias vertientes para obtener una estimación de la función de onda. Una de ellas es modelar directamente a la función de onda. En un inicio se propuso a esta función como el producto de orbitales, un orbital por cada electrón. Sin embargo, se sabe que la función de onda verdadera debe de cambiar de signo cuando las coordenadas entre dos electrones se intercambian, se le conoce a este hecho como principio de antisimetría, y es un reflejo del comportamiento que deben de exhibir los electrones.⁹ Con esta restricción en 1929 John C. Slater propuso como expresión general de la función de onda a un determinante constituido por orbitales, en vez de un simple producto entre ellos.¹¹ En términos esquemáticos tenemos que

$$\Psi(\mathbf{r}_1, w_1, \dots, \mathbf{r}_N, w_N) = \det[\chi_1(\mathbf{r}_1, w_1), \chi_2(\mathbf{r}_2, w_2), \dots, \chi_N(\mathbf{r}_N, w_N)] \quad (2)$$

En esta ecuación hemos remarcado el hecho de que el electrón 1 tiene como coordenadas a $r_1 = x_1, y_1, z_1$, por la parte espacial, y una variable de espín, w_1 . Para N electrones tendremos 4N coordenadas. Por cada electrón se propone dentro del determinante una función, un orbital, y por lo tanto el determinante será de tamaño $N \times N$. El problema ahora es determinar los orbitales, las funciones, que deben de ser incluidos en el determinante. Afortunadamente la función de onda tiene asociado un principio variacional, donde se sabe que la función de onda del sistema será aquella que minimice a la energía, en otras palabras, la función de onda debe de minimizar a

$$E = \int \dots \int d\mathbf{r}_1 d\mathbf{w}_1 \dots d\mathbf{r}_N d\mathbf{w}_N \Psi^*(\mathbf{r}_1, \mathbf{w}_1, \dots, \mathbf{r}_N, \mathbf{w}_N) \hat{H} \Psi(\mathbf{r}_1, \mathbf{w}_1, \dots, \mathbf{r}_N, \mathbf{w}_N) \quad (3)$$

Como en la ecuación 2 se ha escrito a la función de onda en términos de orbitales entonces la minimización de la energía debe de ser con respecto a los orbitales. Cuando tenemos en cálculo diferencial el problema de minimización con respecto a una variable resulta que para encontrar el mínimo, después de igualar la derivada a cero, debemos de encontrar las raíces de una función, en el mejor de los casos un polinomio. En el caso de la minimización de la ecuación 3 lo que obtenemos es una ecuación diferencial en lugar del polinomio y en lugar de la búsqueda de raíces se busca a las funciones que satisfacen la ecuación diferencial con

$$\hat{f} \chi_i(\mathbf{r}w) = \varepsilon_i \chi_i(\mathbf{r}w) \text{ con } i=1, 2, \dots, N \quad (4)$$

Podemos darnos cuenta que esta ecuación se parece mucho a la ecuación 1. Sin embargo, existen dos diferencias fundamentales entre ellas. 1) La ecuación 4 está asociada a un orbital, así se ha transformado el problema de resolver una ecuación que depende de N partículas a N problemas que dependen de las coordenadas de una partícula. 2) El operador \hat{H} de la ecuación 1 se conoce exactamente al saber las interacciones que existen dentro de una molécula lo cual no es cierto para el operador \hat{f} ya que este operador contiene partes que se escriben en términos de los orbitales. ¿Cómo encontrar a los orbitales si no conocemos el operador con el que los vamos a obtener? La forma en que se procede para resolver este problema es proponer un operador de prueba, con este operador de inicio se obtiene un conjunto de orbitales, con estos orbitales construimos nuevamente al operador \hat{f} y así sucesivamente. Este proceso iterativo se termina cuando decidimos que la energía del sistema es lo suficientemente parecida a la iteración previa o que el potencial entre una iteración y otra no cambia sustancialmente. Hemos descrito muy brevemente el método conocido como Hartree-Fock, el cual es un pilar en los métodos de la química cuántica computacional hoy en día, tanto por sus ideas como por la manera en que se resuelve la ecuación 4.

La solución de la ecuación 4 es todo un reto, primeramente lo que se hace es imponer la variable de espín en cada orbital. Esto es muy importante ya que esto determina el espín total del sistema. Para el espín solamente tenemos dos posibilidades ya que puede

tener el valor de $\frac{1}{2}$ o $-\frac{1}{2}$ y esto nos lleva a poder escribir a cada orbital como el producto de dos funciones, una que depende de la parte espacial, φ , y otra de la parte de espín, σ , esto es

$$\chi_i(\mathbf{r}, w) = \varphi_i(\mathbf{r})\sigma_i(w) \quad (5)$$

Evidentemente, dependiendo del espín que tenga asociado cada orbita el sistema podrá tener un espín neto de 0, $\frac{1}{2}$, 1, $\frac{3}{2}$, ... Lo cual define si el sistema se encuentra en un sigulete, doblete, triplete, etc.

La necesidad de una computadora

Para la solución de la ecuación 4, teniendo en cuenta la ecuación 5, Clemens C. J. Rootaan¹² y George G. Hall¹³ hicieron de manera independiente una propuesta muy sencilla para poder ser implementada en sistemas moleculares, basándose en el método de Ritz. Es importante mencionar que este proceder es lo que hoy domina el mundo de la química cuántica computacional. La idea es muy sencilla ya que cada orbital se representa en términos de un conjunto de funciones conocidas. Supongamos que se propone al conjunto de funciones $\{f_k(\mathbf{r})$ con $k = 1, 2, \dots, K\}$, entonces cada orbital i será representado por la combinación lineal de estas funciones de base de la siguiente manera

$$\varphi_i(\mathbf{r}) = \sum_{k=1}^K c_k^i f_k(\mathbf{r}). \quad (6)$$

Como las funciones f_k son conocidas entonces lo que ahora se necesita buscar es el conjunto de coeficientes c_k^i . Nótese que basta un conjunto de funciones de base ya que el conjunto de c_k^i será diferente para cada orbital. Al sustituir la ecuación 6 en 5 y ésta en 4 se obtiene un problema algebraico para los coeficientes c_k^i con la siguiente forma

$$\mathbf{FC} = \epsilon \mathbf{SC} \quad (7)$$

En esta ecuación cada una de las cantidades representa una matriz, un arreglo de $K \times K$ elementos. Al tener un operador de inicio, en nuestro proceso iterativo, la matriz F se conoce y tiene varias contribuciones, todos sus elementos son el resultado de evaluar integrales: la contribución de la energía cinética de los electrones, la interacción electrostática núcleo-electrón y electrón-electrón. Además aparece un término adicional que surge como consecuencia de escribir a la función de onda como un determinante, para que la función de onda sea antisimétrica. A este término adicional a la energía se le conoce como energía de intercambio. La matriz S está compuesta de integrales del producto entre las mismas funciones de base

$$S_{ij} = \int d\mathbf{r} f_i(\mathbf{r})f_j(\mathbf{r}). \quad (8)$$

Para la construcción de una parte de la matriz F es necesario evaluar integrales que se conocen como integrales bielectrónicas y que son de la forma

$$\iint d\mathbf{r}d\mathbf{r}' \frac{f_i(\mathbf{r})f_j(\mathbf{r})f_k(\mathbf{r}')f_l(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (9)$$

Son precisamente estas integrales las que son un cuello de botella en el uso del método de Hartree-Fock. Pensemos en que usamos 10 funciones de base para estudiar a la molécula de agua con el método de Hartree-Fock. En este caso se tendría que evaluar 100 integrales (10×10) para la matriz S. Claro, se puede ver que hay simetría en algunas integrales y no se tendrán que evaluar todas ya que $S_{ij} = S_{ji}$. En realidad cuando se tiene integrales entre dos funciones de base no se tendrá problema alguno. Sin embargo, pensemos en el mismo caso para las integrales bielectrónicas, con ellas se tendrán que evaluar 10000 integrales ($10 \times 10 \times 10 \times 10$). También es cierto que para estas integrales se tiene simetría y hay que evaluar del orden $K^4/8$ integrales. En estos días, para el estudio de un sistema de interés farmacéutico como la β -ciclodextrina se requieren del orden de 1800 funciones de base. Eso significa que tendremos matrices de un tamaño de 1800×1800 (3240000 elementos) y alrededor de 1.3122×10^{12} integrales bielectrónicas que evaluar. De aquí es claro que es necesario el uso de una computadora.

Computadora o supercomputadora

No hay duda que se requiere de una computadora para poder hacer uso del método de Hartree-Fock. Esta conclusión fue rápidamente mencionada por J. C. Slater quien fue uno de los impulsores del método de Hartree-Fock ya que tanto Slater como los miembros de su grupo tenían a la mano calculadoras de mano en la década de los 30. Y no fue que hasta finales de la decada

de los 50 e inicios de los 60 tuvieron la disponibilidad de usar una computadora IBM 790. Precisamente el ejemplo que hemos usado de la molécula de agua se estudió con un método aproximado al de Hartree-Fock (conocido como el método X_a) a inicios de los 60. A diferencia de lo que podemos encontrar en este año 2010, cuando comenzó la química cuántica computacional era muy difícil tener disponibilidad de una computadora. Solamente las grandes universidades o los grandes consorcios podían acceder a una computadora. Existen historias muy interesantes o chuscas de cómo se hacía la química cuántica computacional. Por ejemplo, se sabe que el papá de Douglas R. Hartree le ayudó a evaluar integrales para realizar el método iterativo. No todos contamos con papás que nos ayuden en nuestro trabajo. También se cuenta que la Universidad Autónoma Metropolitana (UAM) contaba con una computadora en el edificio de la rectoría general, en ese entonces se encontraba en el norte de la ciudad de México. Cada investigador en su respectiva unidad (Azcapotzalco, Iztapalapa o Xochimilco) tenía que perforar las tarjetas donde se tenía las instrucciones para correr un programa y después un mensajero en su motocicleta las llevaba a las oficinas de la rectoría general. Pues resulta que en una ocasión el mensajero sufrió un accidente automovilístico y consecuentemente las tarjetas salieron volado. Afortunadamente el mensajero salió ileso, además un buen samaritano acomodó las tarjetas ya que el mensajero se había ido en una ambulancia. Evidentemente, el nuevo arreglo de las tarjetas perforadas no tenía ya ningún sentido. Otra historia menos chusca se dio en la ciudad de Turín, Italia. Donde en los años 70 el grupo de química teórica carecía de una computadora potente para hacer uso del programa que estaban desarrollando para aplicar el método de Hartree-Fock en sólidos. Resulta que una de las compañías televisivas con presencia importante en ese país contaba con una máquina potente y le permitió la entrada a los investigadores del grupo de química teórica para hacer sus cálculos, pero con una condición; deberían de usar la computadora solamente por las madrugadas, así que por varias semanas los investigadores preparaban su programa y sus corridas para ser ejecutadas por las madrugadas, naturalmente no ha de haber sido nada gracioso para los investigadores ir a hacer corridas de su código por las madrugadas, en particular en la época de invierno.

En la década de los 80 y 90 el uso de la computadora se amplió rápidamente a más secciones de la población y consecuentemente las grandes universidades y consorcios importantes adquirieron computadoras con una gran potencia. De hecho en 1993 se creó el sitio <http://www.top500.org> donde se tiene una lista de las 500 computadoras más rápidas del mundo y es precisamente donde la palabra supercomputadora es resaltada. Podemos decir que una supercomputadora se distingue por su potencia al ser comparada con las computadoras contemporáneas a ella. En 1994 la UAM colocó en esta lista una supercomputadora con una velocidad de 3.72 Gflops (*billions of floating point operations per second*, billones de operaciones de punto flotante por segundo) y en el 2008 la misma universidad colocó en esa lista una computadora con una velocidad de 18480.00 Gflops. Como se ve es una diferencia importante en los 14 años que transcurrieron entre la publicación de una lista a otra. En este momento la realización de este artículo se está llevando a cabo en una computadora personal que consta de dos cores, su potencia es por mucho mayor a la de las primeras computadoras hechas con transistores de los años 70. Sin embargo, la supercomputadora instalada en la UAM en el 2008 consta de 2160 cores, tres ordenes de magnitud mayor a la mencionada computadora personal.

El hardware de ahora

Los primeros códigos que resolvían las ecuaciones de Hartree-Fock estaban basados en computadoras seriales, esto significa que todo el trabajo lo hacía un procesador. Muchos de estos códigos prevalecieron por que los procesadores incrementaron su velocidad de manera importante, pero desafortunadamente para esos códigos el incremento de la velocidad de un procesador a llevado a un sobrecalentamiento del mismo. Por este motivo las compañías fabricantes de procesadores han decidido no incrementar la velocidad del procesador, pero a cambio han dividido al procesador en cores. En este momento podemos encontrar en los supermercados computadoras personal con dos cores o en universidades o empresas servidores con procesadores de 4 cores.

Actualmente la tecnología ha permitido tener servidores que gastan eficientemente la energía eléctrica además de construirlos en un espacio reducido, contrario a lo que se tenía hace algunos años. Además, los costos se han reducido de manera apreciable y han permitido que se puedan construir supercomputadoras a partir de la unión de varios servidores. Lo que ha llevado a la construcción de los ahora llamados cúmulos o "clusters" de servidores y que están predominando en el área del supercómputo. Sin embargo, se avecina una nueva generación de supercomputadoras. Es bien sabido que la industria de los videojuegos ha invertido bastante para tener consolas que hagan un despliegue gráfico eficiente y que contengan el mayor número de efectos visuales. De manera natural han trabajado con tareas en paralelo, además las tarjetas de video han sido diseñadas para ser autónomas del procesador ya que tienen sus propios circuitos y su propia memoria RAM. En el 2001 se empezó a ver la posibilidad de llevar a

cabo aplicaciones científicas sobre tarjetas gráficas, las cuales se conocen como GPUs (*Graphics Processing Units*, unidades de procesamiento gráfico). Para el 2006 ya se vio como una realidad el hacer cálculos científicos sobre los GPUs. Evidentemente esto lleva a nuevas técnicas de programación para usar de manera simultánea a los CPUs y GPUs en un ambiente de clusters.

Nuevas técnicas de programación

La actual tecnología demanda programar con otra filosofía a la programación serial. De hecho es ahora común que un programa de química cuántica computacional esté basado en MPI (*message passing interface*, interface de paso de mensaje) el cual ha sido un estandar de programación desde 1993.¹⁴ ¿Cuánto se puede ganar al programar en paralelo? En principio, un programa diseñado para correr en paralelo debe de ejecutar tareas en menor tiempo cuando se usan más procesadores. Pero no sólo eso, también el ejecutar un programa en paralelo significa que se puede incrementar la memoria RAM en la medida de que se usen varios procesadores en esa ejecución. Ya hemos mencionado a la β -ciclodextrina, la cual ha sido estudiada usando 1806 funciones de base. En este momento usando códigos diseñados para correr en paralelo, como el NWChem,¹⁵ es factible obtener los mínimos de este sistema en la superficie de energía potencial. Lo cual permite predecir las conformaciones espaciales que puede adoptar esta molécula. Con el arribo de los GPUs será necesario tener un estandar de programación, como el MPI, ya que en este momento hay dos propuestas claras que son independientes una de otra.¹⁶

Comentarios finales

A lo largo de este documento se ha hecho referencia al método de Hartree-Fock, sin embargo es un método que se usa como punto de partida de métodos más sofisticados basados en la función de onda conocidos como métodos correlacionados.^{9,17}

Adicionalmente a los métodos basados en la función de onda, existe la Teoría de Funcionales de la Densidad¹⁸ (TFD) la cual ha sido exitosa en los últimos años ya que por su simplicidad ha permitido tener implementaciones computacionales que permiten estudiar sistemas electrónicos que cuentan con un gran número de átomos, del orden de cientos. Sin embargo, muchas de las ideas de la implementación del método de Hartree-Fock son vigentes en la implementación de la TFD. Es importante mencionar que esta teoría domina actualmente las aplicaciones de la química cuántica computacional y será vigente por varios años por las bondades que ofrece. De hecho la TFD ya ha sido implementada en computadoras que usan de manera indistinta los GPUs y CPUs.¹⁹ Tal vez, en un futuro cercano, podamos ver aplicaciones de la química cuántica que sean ejecutadas en consolas de video juego.

Referencias

- 1.- Schrodinger, E. *Ann. d. Physik* (1926) 79, 361.
- 2.- Ibid 79, 489.
- 3.- Ibid 79, 734.
- 4.- Ibid 80, 437.
- 5.- Ibid 81, 109.
- 6.- Schrodinger, E. *Die Naturwissenschaften* (1926) 14, 664.
- 7.- Schrodinger, E. *Phys. Rev.* (1926) 28, 1049
- 8.- Landau, L. D. y Lifshitz, E. M. *Mechanics*, Third Edition: Volume 1 (Course of Theoretical Physics). Butterworth-Heinemann, 1976.

- 9.- Levine, I. N. *Quantum Chemistry*, 6ta. Economy Edition. Prentice Hall India 2008.
- 10.- Born, M. y Oppenheimer, J. R. *Ann. D. Physik* (1927) 84, 457.
- 11.- Slater, J. C. *Phys. Rev.* (1929) 34, 1293.
- 12.- Roothaan, C. C. J. *Rev. Mod. Phys.* (1951) 23, 69.
- 13.- Hall, G. G. *Proc. Royal Soc. London* (1951) A 205, 541.
- 14.- Sitio de web <http://www.mcs.anl.gov/research/projects/mpi>.
- 15.- Sitio de web <http://www.emsl.pnl.gov/capabilities/computing/nwchem>
- 16.- Kirk, D. B. y Hwu, W. W. *Programming massively parallel processors. A hands-on approach*. Morgan Kaufmann y Elsevier (2010).
- 17.- Szabo A. y Ostlund N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory. Second Edition revised*. Dover Publications, 1996.
- 18.- Parr, R. G. y Yang, W. *Density Functional Theory Of Atoms And Molecules*. Oxford University Press, 1989.
- 19.- Genovese, L.; Ospici, M.; Deutsch, T.; Méhaut, J.-F.; Neelov, A. y Goedecker, S. J. *Chem. Phys.* (2009) 131, 034103.