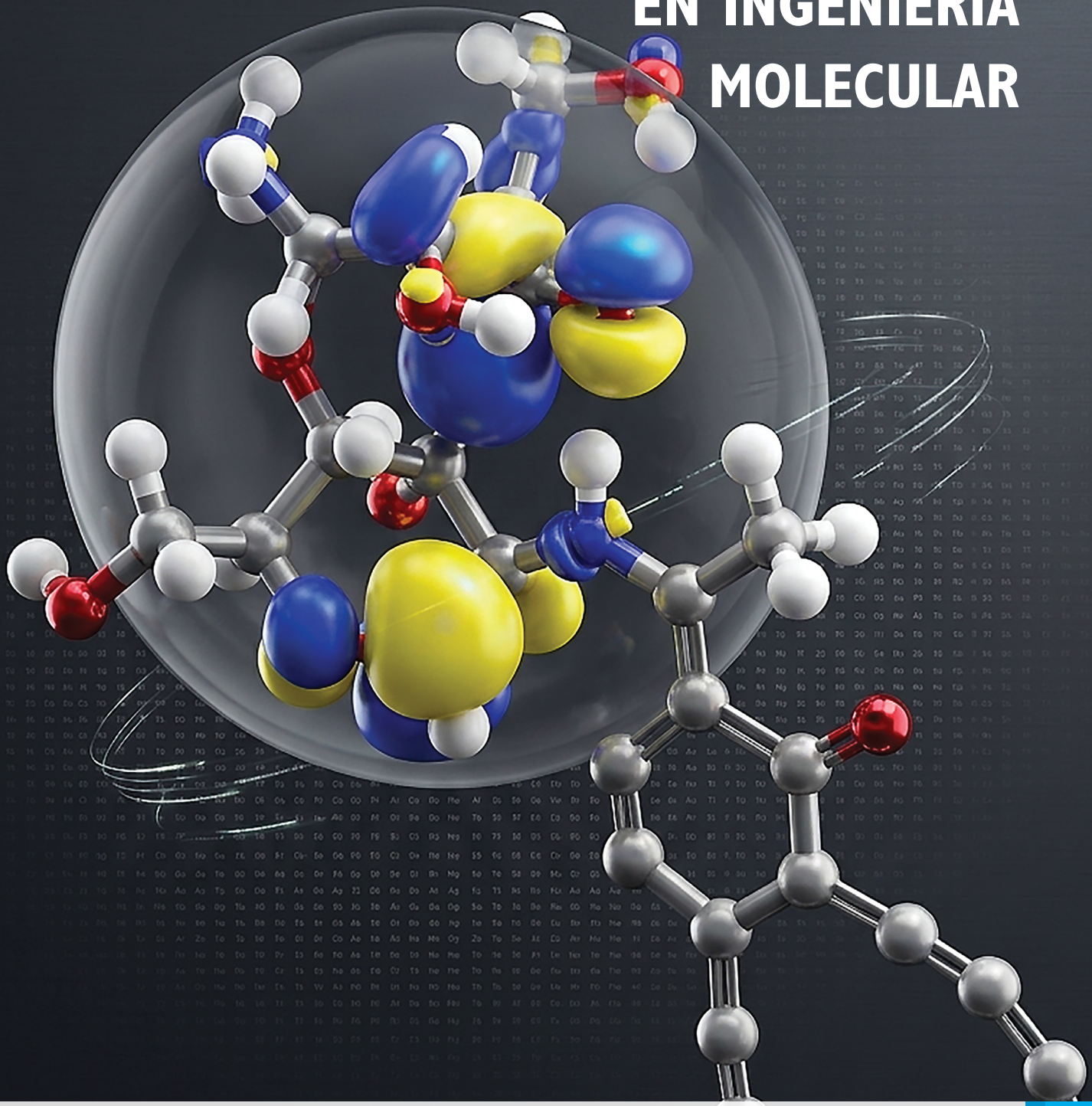
FUNDAMENTOS DE FÍSICA EN INGENIERÍA MOLECULAR



EDICIONES COMUNICACIÓN CIENTÍFICA

Juan Horacio Pacheco Sánchez

Fundamentos de física en ingeniería molecular

EDICIONES
COMUNICACIÓN
CIENTÍFICA

Ediciones Comunicación Científica se especializa en la publicación de conocimiento científico de calidad en español e inglés en soporte de libro impreso y digital en las áreas de humanidades, ciencias sociales y ciencias exactas. Guía su criterio de publicación cumpliendo con las prácticas internacionales: dictaminación de pares ciegos externos, autentificación antiplagio, comités y ética editorial, acceso abierto, métricas, campaña de promoción, distribución impresa y digital, transparencia editorial e indexación internacional.

Cada libro de la Colección Ciencia e Investigación es evaluado para su publicación mediante el sistema de dictaminación de pares externos y autentificación antiplagio. Invitamos a ver el proceso de dictaminación transparentado, así como la consulta del libro en Acceso Abierto.

www.comunicacion-cientifica.com

DOI.ORG/10.52501/cc.421



COMUNICACIÓN CIENTÍFICA PUBLICACIONES ARBITRADAS
HUMANIDADES, SOCIALES Y CIENCIAS

CC+I
COLECCIÓN CIENCIA e INVESTIGACIÓN

Fundamentos de física en ingeniería molecular

Juan Horacio Pacheco Sánchez

EDICIONES
COMUNICACIÓN
CIENTÍFICA

1.-

Pacheco Sánchez, Juan Horacio,

Fundamentos de física en ingeniería molecular / Juan Horacio Pacheco Sánchez. — Ciudad de México : Comunicación Científica, 2026.

377 páginas ; 21.59 x 27.94 centímetros

DOI: 1052501/cc.421

ISBN: 978-968-9738-97-8

1. Física molecular. 2. Física. 3. Dinámica molecular.

LC: QC170 P33 DEWEY: 539.7

La titularidad de los derechos patrimoniales y morales de esta obra pertenece al autor D. R. © Juan Horacio Pacheco Sánchez, 2026. Reservados todos los derechos conforme a la Ley. Su uso se rige por una licencia Creative Commons BY-NC-ND 4.0 Internacional, https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode.es

Primera edición, 2026

Diseño de portada: Francisco Zeledón • Interiores: Guillermo Huerta

Ediciones Comunicación Científica, S. A. de C. V., 2026,
Av. Insurgentes Sur 1602, piso 4, suite 400,
Crédito Constructor, Benito Juárez, 03940, Ciudad de México,
Tel.: (52) 55-5696-6541 • Móvil: (52) 55-4516-2170
info@comunicacion-cientifica.com • www.comunicacion-cientifica.com
comunicacioncientificapublicaciones @ComunidadCient2

ISBN 978-968-9738-97-8

DOI 10.52501/cc.421

**Esta obra fue dictaminada mediante el sistema de pares ciegos externos.
El proceso transparentado puede consultarse, así como el libro en acceso abierto, en
https://doi.org/10.52501/cc.421**

a

Elpidio Chacón Esponda†

Mi más profundo reconocimiento a Sidonio Castillo Ánimas† por compartir sus conocimientos en física atómica y molecular conmigo, y por su amistad fraterna. Gracias a él colaboré con Octavio Novaro Peñalosa†, a quien Elpidio Chacón Esponda† me recomendó acercarme. Actualmente, investigar en curvas de potencial como aprendí de Sidonio es fundamental para profundizar en energías renovables, no obstante, por la esencia de lo que aprendí con Elpidio y por su amistad es a quien reconozco como mi Padre académico.

Agradecimientos

A Dios por la gracia que me concede.

A mi Familia por todo su amor.

A todos aquellos que de una u otra manera han contribuido a la realización de este libro. Una mención especial a mi amigo Armando Bravo Ortega por sus atinados comentarios a una primera versión de este libro.

Finalmente, al Tecnológico Nacional de México / Instituto Tecnológico de Toluca / División de Estudios de Posgrado e Investigación por las facilidades otorgadas para construir un soporte a Ingeniería Molecular.

comunicacion-cientifica.com

Índice

Introducción 17

Capítulo 1. Estructura molecular y potenciales interatómicos 23

Teoría de Bohr de los términos de Balmer 23

Ecuación de Schrödinger 25

La aproximación de Born-Oppenheimer 28

Teoría orbital molecular 36

Sistema de dos partículas 36

Degeneración 66

Momento angular 66

Método LCAO 72

Orbitales de moléculas diatómicas 76

Orbitales de moléculas poliatómicas 80

Hibridación de los orbitales atómicos 81

Hibridación $\boldsymbol{sp3}$ 81

Hibridación $s\boldsymbol{p2}$ 82

Hibridación sp^1 83

Simetrías 85

Operaciones 86

Fuerzas intermoleculares 90

Introducción 90

Fuerzas de Coulomb 92

Interacción ion-ion 92

Interacción ion-dipolo 93
Interacción dipolo-dipolo 96
Fuerzas de inducción 100
Interacción ion-no-polar 101
Interacción dipolo-no-polar (energía de Debye) 103
Fuerzas de dispersión de London o de Van der Waals 105
La ecuación de London 106
Efecto de retardamiento 119
Efectos del solvente 125
Interacciones moleculares electrostáticas 127
Puentes de hidrógeno 129
Interacción hidrofóbica 130
Interacciones hidrofílicas 131
Potenciales de interacción 133
Un modelo de potencial 134
Potenciales intermoleculares 136
Cálculos mecánico-cuánticos de fuerzas intermoleculares 138
Interacción átomo-molécula 139
Interacción molécula-molécula 141

Capítulo 2. Métodos de primeros principios 143
Introducción 143
Método de campo autoconsistente 144
Orbitales espaciales y espín-orbitales 150
Principio de exclusión de Pauli y determinante de Slater 156
Ecuaciones de Hartree-Fock 157

Variación de E 164
Solución de Roothaan a ecuación H-F 172
Funciones base 177
Densidad electrónica 179
Corrección electrónica 182
Correlación electrónica 182
Configuración de funciones de estado 183
Interacción de configuraciones 184
Métodos de multiconfiguración y multireferencia 191
Teoría de perturbación de muchos cuerpos de Möller-Plesset 193
Teoría de funcionales de la densidad 196
La idea original del modelo de Thomas-Fermi 196
Partícula en una caja 197
Energía como un funcional de la densidad 203
Teoremas de Hohenberg y Kohn 203
Ecuación de Kohn-Sham 204
Potencial químico 206
Métodos del gradiente y propiedades moleculares 213
Optimización de geometría 214
Sistemas de gradiente dinámico 215
Métodos semiempíricos 223
Método MO de electrón-libre (FEMO) 224
Método MO de Huckel (HMO) 225
Método Pariser-Parr-Pople (PPP) 225
Mecánica molecular 226

Introducción a campos de fuerza (*force fields*) 226
COMPASS como modelo de campo de fuerza (*force field*) 230
Modelo covalente 231
Modelo iónico 232
Modelo semiiónico para óxidos metálicos 233
Modelo de Lennard-Jones para metales 234

Capítulo 3. Dinámica molecular 235
Ecuaciones de movimiento para sistemas atómicos 235
Tres modelos concordantes 235
Formulación newtoniana 237
Formulación lagrangiana 237
Formulación hamiltoniana 238
Conservación de la energía 239
Energía potencial 240
Métodos de diferencias finitas 240
Teorema de Taylor 240
Algoritmo de Verlet 242
Esquema salto de rana (*leap-frog*) 244
Dinámica molecular de cuerpos rígidos no-esféricos 245
Moléculas no-lineales 252
Precisión 255
Dinámica molecular de esferas duras 256

Capítulo 4. Modelación y simulación con método de Montecarlo 263
Integración Montecarlo 273

Ejemplo numérico 275
Importancia del muestreo 280
El método de Metropolis 290
Montecarlo isotérmico-isobárico 294
Líquidos moleculares 298
Moléculas rígidas 298
Moléculas no-rígidas 301

Capítulo 5. Uso de software de simulación molecular 307
Metodología de una dinámica molecular 308
Propiedades macroscópicas 311
Características de GROMACS 312
Generación de archivos para GROMACS 312
Función de distribución radial (RDF) 315
Disponibilidad del programa GROMACS 315
Paquetes de *software* para cálculos de estructura electrónica 316
Cálculos *ab initio* con la ecuación de Schrödinger 316
Lista de programas computacionales 317
Rol de la computación en la química 318
Programas acoplados secuencialmente 319
Procedimiento de cálculos ab initio PSHONDO-IJKL-GMCP-FOCK-CIPSI 324
Pseudopotenciales 332
Disponibilidad de programas 337

Anexos

Anexo A. Relación fundamental de la termodinámica 341

Anexo B. Energía de interacción de sistemas dipolares 347

Anexo C. Teoría de perturbaciones estacionarias 349

Anexo D. Átomo de hidrógeno 353

Anexo E. Expansión multipolar 361

Anexo F. Gradiente del potencial 365

Anexo G. Quaterniones 367

Anexo H. Trompo 369

Referencias 371

Sobre el autor 377

comunicacion-cientifica.com

Resumen

En esta obra se abordan los conceptos y fundamentos de física en los cuales se soporta la ingeniería molecular. Se espera que cada uno de los capítulos sea oportuno, apropiado y relevante para el lector, con el fin de que sea introducido a este tema y se sienta estimulado a explorar y buscar respuestas a través de la simulación molecular.

El libro contiene desarrollos algebraicos para que el lector comprenda la esencia de cálculos ocultos cuando se ejecuta un programa de simulación molecular con el enfoque que aquí se presenta, especialmente pensados para quienes no han tenido un contacto profundo con cálculos computacionales.

Se lleva una secuencia de lectura con figuras ilustrativas, por lo que se omitió, hasta donde fue posible, el uso de numeración de ecuaciones, resaltando solo aquellas que se vuelven a mencionar en alguna otra parte del texto. Además, se escribieron anexos en algunas partes para no extender demasiado el texto y no perder la continuidad en la lectura.

Palabras clave: *ingeniería molecular, fundamentos de física, energías potenciales interatómicas, dinámica molecular*.

Introducción

Este libro lo diseñé y escribí con la intención de tender un puente que permita a los egresados de las carreras de ingeniería introducirse en el ambiente de la simulación molecular. Esto ocurre frente al creciente interés que ha propiciado el modelamiento molecular, el cual requiere de cierta formación académica en modelos teóricos fisicoquímicos usados para realizar simulaciones moleculares y posibilita resolver problemas que corresponden directamente a la interacción entre átomos y moléculas. Esta gran inquietud ha surgido con el crecimiento computacional y sus facilidades para plantear y encontrar algún modelo que ayude a resolver problemas reales, tales como aquellos que ocurren al buscar energías renovables o no renovables, construir nuevos materiales, disminuir la contaminación ambiental, atacar enfermedades en seres vivos y muchos más.

Buena parte de este manuscrito lo expuse previamente en aulas durante algunos años, y ha servido para que más de una decena de alumnos obtenga su grado. Entre los trabajos desarrollados se encuentran los relacionados con el diseño de electrodos de celdas de combustible tipo PEM (Pacheco-Sánchez y Mora-Zarco, 2017; Pacheco-Sánchez et al., 2019; Valenzuela et al., 2019), la porosidad de carbón activado útil en pilas entre otras aplicaciones (Esquivel-Aguilar, 2015; Pacheco-Sánchez et al., 2017; Rivas, 2016), el almacenamiento de hidrógeno (Desales-Guzmán et al., 2018; Isidro-Ortega et al., 2017, 2018; Zaragoza et al., 2014), el tratamiento de pesticidas (Esquivel-Aguilar et al., 2019), la limpieza del agua (Suárez y Pacheco, 2017; Suárez-Reyes et al., 2018; Pacheco-Sánchez, 2019a), las interacciones fármaco-receptor (Flores-Ramírez y Pacheco-Sánchez, 2019), el fármaco desintoxicante (Hernández-Benites y Pacheco-Sánchez, 2018; Pacheco-Sánchez, 2019b), la energía asociada a una monocelda de combustible (Larrea, 2017) y estados de transición en síntesis (Márquez et al., 2016).

El hidrógeno como energía renovable tiene el alto potencial de ser portador de energía ideal: fuente de energía segura, limpia y no contaminante, pero su uso es limitado debido a su aspecto de almacenamiento, que requiere condiciones de alta presión y/o baja temperatura. Otro método de almacenamiento que no utiliza alta presurización necesita que el hidrógeno sea inyectado en un material poroso con una gran área superficial y alta porosidad. Los materiales porosos de zeolita, carbono y la forma modificada de ambos materiales

denominada *zeolite templated carbon* (ZTC) se están considerando para la aplicación de almacenamiento de hidrógeno. La infiltración de precursores de carbono en el poro de zeolita, seguida de la eliminación de la plantilla de zeolita en la producción de ZTC, da como resultado materiales que tienen poros tridimensionales de alta regularidad con una gran superficie específica. En algunos estudios experimentales se puede ver que la aplicación de almacenamiento de hidrógeno en ZTC se considera prometedora. Mas aún, los cálculos teóricos se han acercado mucho al diseño de estos materiales almacenadores de hidrógeno, que ya se intentan producir en laboratorios.

Así mismo, los estudios teóricos han ido adquiriendo cada vez mayor relevancia. Por ejemplo, la catálisis era una ciencia semiempírica hasta hace no mucho tiempo. Por su parte, la química cuántica dentro de los fenómenos catalíticos da una estimación i) de los fenómenos de adsorción de moléculas de gas sobre superficies metálicas, ii) de las energías de activación y iii) de los mecanismos de reacción de las especies adsorbidas. La tendencia hacia una estrecha relación entre las investigaciones teóricas y experimentales que se realizan actualmente en esta área hace atractivo e importante el estudio teórico de la reactividad y la naturaleza de la interacción metal-gas.

El estudio del rompimiento C-H en la molécula de metano presenta una ruta alterna para obtener hidrocarburos saturados y olefinas. La vinculación entre lo experimental y lo teórico ha ocurrido a través de diferentes técnicas experimentales, como la caracterización espectroscópica, el estudio de fotoexcitación, los rayos X, etc. Estas técnicas dan información asociada a la reactividad respecto a los diferentes estados electrónicos que presentan los metales en un proceso. El estudio de la interacción metal-hidrógeno molecular ha permitido comparar los resultados obtenidos experimentalmente con los correspondientes predichos por química cuántica. El caso de la molécula de metano depende de la naturaleza del metal empleado. Más aún, los cálculos teóricos de sistemas en los que participan metales de transición se dificultan debido a problemas inherentes a la correlación electrónica, así como a la gran cantidad de estados electrónicos que surgen como consecuencia de la consideración de la capa ‘d’ metálica no-llena.

Acerca de la limpieza del agua, los colorantes que están presentes en las aguas residuales, especialmente los de origen sintético, son responsables de numerosos efectos nocivos para el medioambiente, la flora y la fauna acuática. Entre los efectos más importantes

se encuentran la reducción del oxígeno disuelto, la eutrofización, la formación de compuestos recalcitrantes y tóxicos en las células y la obstrucción del paso de la luz a los cuerpos de agua y el deterioro estético. Por ejemplo, el rojo 2 es un colorante azo que puede causar intolerancia en las personas afectadas por los salicilatos. Además, es un liberador de histamina y podría intensificar los síntomas del asma. También está involucrado en casos de hiperactividad en niños cuando se usa en combinación con benzoatos. Muchos países, entre ellos Estados Unidos, han limitado su uso porque crea problemas que requieren tratamiento médico. Cuando los colorantes de tipo azo, también llamados reactivos, se vierten en los cuerpos de agua producen aminas como consecuencia de la ruptura del enlace azo, lo cual provoca muchos efectos en órganos humanos como el cerebro, hígado, riñón, sistema nervioso central y reproductivo.

La interacción entre quitosano (adsorbente) y un colorante azoico (adsorbato), a través de sus sitios de reacción (ion de amoniaco del grupo amino protonado y ion trióxido de azufre del grupo sulfonato), respectivamente, proporciona el tipo de adsorción existente. Esto se hace primero calculando teóricamente su optimización geométrica y luego construyendo la superficie de energía potencial fijando el ion de amoniaco y trayendo el ion trióxido de azufre. La reacción entre estos iones moleculares produce ácido sulfámico que en iones híbridos forma una molécula más estable que la forma ácida neutra en estado sólido. Las propiedades físicas (estructurales y espectroscópicas) y las reacciones químicas del ácido sulfámico son bien conocidas. La estructura de la forma de iones híbridos y de la neutra se ha analizado teóricamente con el uso de cálculos computacionales. El caso neutro se ha estudiado como isómero, y se conoce el uso de amoniaco para eliminar el dióxido de azufre.

La interacción entre un antibiótico (penicilina G) y un receptor (e-coli) es posible mediante cálculos de superficie de energía potencial (PES) y dinámica molecular. Estos cálculos están relacionados con la estructura de las moléculas. Se realizan cálculos paso a paso de un solo punto de la interacción de penicilina G con moléculas de e-coli a diversas conformaciones geométricas para obtener pozos de potencial, con el fin de saber si hay atracción o repulsión entre ellos. Al calcular y graficar, las energías mínimas de los pozos indican la posible existencia de fisisorción entre estas moléculas. Además, a partir de una dinámica molecular de sistemas de moléculas de penicilina G con e-coli, se logra calcular la distribución radial para obtener la dispersión de las moléculas según la distancia; entre otras

cosas, también es indicativa de las interacciones entre las moléculas. Con estos cálculos es posible inferir si el microorganismo e-coli se puede extraer mediante fisisorción con antibiótico penicilina G.

La ingeniería molecular ha sido básica en la construcción de resultados de modelos que, por un lado, ayudan al diseño de síntesis experimentales que son muy costosos y, por otro lado, convalidan mediciones experimentales cuando sus resultados son comparables a estas. Actualmente se cuenta con herramientas de *software* clave para atacar reactividades entre moléculas, que han sido un paradigma exclusivo de químicos e ingenieros químicos. El desarrollo de la física ha sido el ancla para que ahora exista lo que es bien conocido como química computacional, más allá de la reactividad de Lewis, ya que la exploración electrónica en las reactividades entre átomos y moléculas es directa.

La química computacional tiene sus raíces en cálculos *ab initio* con la ecuación de Schrödinger. El término *ab initio* significa 'desde el principio', y se refiere a cálculos derivados de principios teóricos, sin incluir información experimental. Las aproximaciones usadas corresponden al uso de métodos numéricos, tales como los errores de truncamiento o de redondeo al resolver ecuaciones, el uso de cálculo variacional, etc. El cálculo *ab initio* más simple es el de Hartree Fock (HF), en el cual una primera aproximación es la de campo central. Este es un cálculo variacional, y las energías aproximadas son iguales o más grandes que la energía exacta. Una segunda aproximación es que la función de onda se describe por una forma funcional, la cual se conoce solo para unos cuantos sistemas electrónicos. Las funciones usadas son las combinaciones lineales de orbitales tipo Slater o tipo Gaussiano, abreviados STO y GTO, respectivamente. La función de onda está constituida de combinaciones lineales de orbitales atómicos, y más frecuentemente de combinaciones lineales de funciones base.

Aún existe el problema de modelar con gran cercanía a la realidad construyendo sistemas de partículas que, por su enorme tamaño, son muy costosos computacionalmente; tal es el caso cuando se trabaja con proteínas o con micelas, entre otros ejemplos. Se espera que con los modelos computacionales Qubit se puedan atacar estos problemas sin costos computacionales tan grandes. Ya ha sido un gran avance contar con las CUDA, que han permitido trabajar con tamaños gigantescos de partículas con el *software* GROMACS de Dinámica Molecular.

El lector asiduo de este libro logrará introducirse y adquirir fundamentos básicos para entender y entrenarse en las aproximaciones moleculares en modelos útiles en la solución de problemas afines, que se encuentran en el desarrollo de la ciencia y la tecnología, como una alternativa a los desarrollos experimentales, que están fundamentados en las leyes de la física, la química, la biología, las matemáticas, etcétera, y que, además, tienen aplicación en la ingeniería, la medicina, la electrónica de alto rendimiento y otros.

Este libro cuenta con cinco unidades. La unidad 1 trata la estructura molecular y los potenciales interatómicos; entre otras cosas, nos introduce a los postulados de la mecánica cuántica, la teoría de orbitales moleculares y fuerzas intermoleculares. La unidad 2 aborda los métodos de primeros principios e incluye brevemente la teoría de funcionales de la densidad, los métodos semiempíricos y la mecánica molecular. La unidad 3 se refiere a los principios básicos de mecánica clásica y termodinámica necesarios en la dinámica molecular. La unidad 4 contiene una breve introducción al método de Montecarlo. Finalmente, en la unidad 5 se habla tanto del uso del *software* de simulación molecular que ha sido empleado por nuestro grupo, como de uno o dos más que seguimos explorando.

Capítulo 1. Estructura molecular y potenciales interatómicos

La estructura molecular aparece con el descubrimiento de la tabla periódica de los elementos y con la solución teórica al problema de radiación de cuerpo negro desarrollada por Planck (Atkins, 2001; Atkins y Friedman, 1996, 2005; Herzberg, 1944; Paglia, 1971); al analizar uno de los osciladores armónicos que usó, Planck encontró en su solución que la energía es proporcional a la frecuencia. A la constante de proporcionalidad, más conocida como constante de Planck, se le identifica como cuanto de energía. Con la solución al problema del efecto fotoeléctrico, que consiste en el proceso por el cual se liberan electrones de un metal por la acción de la radiación, Einstein demostró que la constante de Planck relaciona la cantidad de energía que transporta un fotón con la frecuencia de su onda electromagnética, y quedó abierta nuevamente la idea de que la luz se comporta como partícula. Previamente, Maxwell sintetizó las ecuaciones de campo electromagnético considerando a la luz como ondas. Posteriormente, Luis de Broglie demostró que la luz tiene la dualidad onda-partícula, es decir, que se comporta como onda y como partícula, y a esto se debe el éxito del modelo del átomo de Bohr. Así, paso a paso va naciendo la mecánica cuántica y, en consecuencia la química cuántica, con un fuerte fundamento en la ecuación de onda-partícula de Schrödinger; es él mismo quien logra describir teóricamente el comportamiento del átomo de hidrógeno. De esta manera se inician estudios de energía potencial de interacción de sistemas atómicos y moleculares.

Teoría de Bohr de los términos de Balmer

Para explicar la característica de emisión de luz por átomos, Bohr propuso dos suposiciones básicas:

i. Del número infinito de órbitas de un electrón alrededor de un núcleo atómico, las cuales son posibles de acuerdo con la mecánica clásica, *solo ciertas órbitas discretas* ocurren realmente. Por tanto, en contradicción a la teoría clásica de Maxwell, el electrón, a pesar del movimiento acelerado, no emite ondas electromagnéticas (luz) mientras esté en una de estas órbitas discretas.

ii. Se emite o adsorbe radiación por una *transición del electrón de un estado cuántico a otro —por un salto cuántico—*; la diferencia de energía entre los dos estados se emite

o absorbe como un cuanto de luz de energía $h\nu'$ ($h = constante\ de\ Planck, \nu' = frecuencia$). La luz cuántica es emitida cuando el átomo va de un estado de energía superior a uno de energía inferior, y es absorbida en el caso inverso (conservación de la energía). Se tiene la relación $h\nu' = E_{n_1} - E_{n_2}$, siendo E_{n_1} y E_{n_2} las energías de los estados superior e inferior, respectivamente. Esta relación es la *condición de frecuencia de Bohr*. El índice n de E distingue las diferentes órbitas y sus valores de energía de una a otra.

Según Bohr, solo ciertas órbitas realmente ocurren y se seleccionan por el postulado de que *el momento angular* $\boldsymbol{L}$ *es un múltiplo entero de* $h/2\pi$, esto es $\boldsymbol{L} = n\frac{\boldsymbol{h}}{\boldsymbol{2\pi}}$. Es decir, como $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ ($\boldsymbol{r}$ es el vector de posición y $\boldsymbol{p}$=$m\boldsymbol{v}$ es el momento lineal «m es masa y $\boldsymbol{v}$ es el vector de velocidad») cuando los ejes son perpendiculares $L = rp\ sen90°$, o bien $L = mvr$, por lo tanto $mvr = n\frac{h}{2\pi}$, donde $n = 1,2,3,\ldots$ es llamado el número cuántico principal. Despejando *r* obtenemos:

$$r = n\frac{h}{2\pi mv} \tag{1}$$

Suponiendo que las órbitas son círculos de radio r, podemos aplicar la segunda ley de Newton. Dado que existen una fuerza que es de atracción coulombiana y una fuerza de aceleración que es la centrípeta, de su balance de equilibrio se tiene que

$$F_{Coulomb} = F_{Centrípeta}$$

$$\frac{Ze^2}{r^2} = \frac{mv^2}{r} \quad o \quad r = \frac{Ze^2}{mv^2} \tag{2}$$

Sustituyendo en la ecuación (1) se determina

$$\frac{Ze^2}{mv^2} = \frac{nh}{2\pi mv} \qquad \rightarrow \qquad v = \frac{2\pi Ze^2}{nh} \qquad \therefore \qquad mv = \frac{2\pi mZe^2}{nh}$$

Al sustituir nuevamente en la ecuación (1) se obtiene

$$r = \frac{nh}{2\pi\frac{2\pi mZe^2}{nh}}$$

Al reacomodar se adquiere el radio de la órbita circular

$$r = \frac{n^2h^2}{4\pi^2me^2Z} \tag{3}$$

para Z=1, y con n=1 se obtiene la órbita más pequeña posible

$$r = \frac{h^2}{4\pi^2 me^2} = a_H = 0.529\text{Å}$$

conocida como radio de Bohr.

Para órbitas circulares, la energía total es

$$E = energía\ potencial + energía\ cinética = -\frac{Ze^2}{r} + \frac{1}{2}mv^2$$

Entonces, usando la fórmula (2) obtenemos

$$E = -\frac{Ze^2}{r} + \frac{1}{2}\frac{Ze^2}{r} = -\frac{Ze^2}{2r}$$

Al sustituir el valor de r de la ecuación (3) y usar μ en lugar de *m*, se obtiene

$$E_n = -\frac{Ze^2}{2\dfrac{n^2h^2}{4\pi^2\mu e^2 Z}} \qquad o\ \ bien \qquad E_n = -\frac{2\pi^2\mu e^4}{h^2}\cdot\frac{Z^2}{n^2} \tag{4}$$

Por la segunda suposición de Bohr $h\nu' = E_{n_1} - E_{n_2}$, la condición de frecuencia es:

$$\nu = \frac{\nu'}{c} = \frac{E_{n_1}}{hc} - \frac{E_{n_2}}{hc} \qquad \Longrightarrow \qquad \nu = \frac{1}{hc}\left(E_{n_1} - E_{n_2}\right)$$

Por lo tanto, los números de onda de las líneas espectrales emitidas están dadas por

$$\nu = \frac{2\pi^2\mu e^4}{ch^3}\cdot Z^2\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right)$$

donde n_1 y n_2 son los números cuánticos principales de los estados superior e inferior.

Balmer midió el espectro del átomo de hidrógeno y propuso la fórmula empírica

$$\nu = \frac{R}{n_2^2} - \frac{R}{n_1^2}$$

para representar dicho espectro. El acuerdo entre las frecuencias resultantes de Bohr (teórico) y de Balmer (experimental) es excelente como punto de partida de la mecánica cuántica.

Ecuación de Schrödinger

Para demostrar que la luz tiene la dualidad onda-partícula, Luis de Broglie propuso que cada objeto con momento también tiene asociada una longitud de onda

$$\lambda = \frac{h}{mv} = \frac{h}{p}$$

La longitud de onda λ de De Broglie es inconmensurablemente pequeña para objetos macroscópicos, tales como una pelota o una bala, que son descritos por la mecánica clásica. Sin embargo, para partículas subatómicas, y en específico el electrón, la longitud de onda de De Broglie es comparable en tamaño a enlaces químicos. Tres años después de que se publicó la hipótesis de De Broglie, Schrödinger definió la mecánica ondulatoria y resolvió el oscilador armónico y el átomo de hidrógeno.

La forma general de una onda puede describirse como una exponencial compleja con amplitud A, frecuencia ω y longitud de onda λ.

$$\psi(x,t) = Ae^{2\pi i\left(\frac{x}{\lambda}-\omega t\right)}$$

Al derivar respecto a *x* unidimensionalmente en el espacio, sustituir la longitud de onda de De Broglie $\lambda = \frac{h}{p}$ y del hecho de que $\hbar = \frac{h}{2\pi}$ se obtiene

$$\frac{d\psi}{dx} = \frac{2\pi i}{\lambda}Ae^{2\pi i\left(\frac{x}{\lambda}-\omega t\right)} = \frac{2\pi i}{\lambda}\psi = \frac{ip}{\hbar}\psi$$

donde $\psi(x,t)$ es una función de onda que depende de la posición y del tiempo, *p* es el momento lineal,

$$\frac{d^2\psi}{dx^2} = \left(\frac{2\pi i}{\lambda}\right)^2 Ae^{2\pi i\left(\frac{x}{\lambda}-\omega t\right)} = \left(\frac{2\pi i}{\lambda}\right)^2\psi = -\frac{p^2}{\hbar^2}\psi = -\frac{2mT}{\hbar^2}\psi$$

$$\Longrightarrow \qquad \frac{d^2\psi}{dx^2} = -\frac{2m(E-V)}{\hbar^2}\psi \qquad o\ bien$$

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V\right]\psi = E\psi$$

Es la ecuación de Schrödinger independiente del tiempo para la cual E es la energía total y H es el operador de energía o hamiltoniano $H = T + V = \frac{p^2}{2m} + V = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V$, para el cual T es la energía cinética, V es la energía potencial y *p* es el momento lineal. Ahora, si derivamos la ecuación de onda respecto al tiempo se tiene que

$$\frac{\partial\psi}{\partial t} = -2\pi i\omega Ae^{2\pi i\left(\frac{x}{\lambda}-\omega t\right)} = -i\frac{E}{\hbar}\psi$$

Para $E = h\omega$. Por lo tanto $E\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t}$, y al sustituir en la ecuación de Schrödinger independiente del tiempo, queda deducida la ecuación de Schrödinger:

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V\right]\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t}$$

Schrödinger encontró una solución analítica para el problema del átomo de hidrógeno. No obstante, con su ecuación se pueden resolver numéricamente sistemas de pocas partículas con simetría bien definida. El cálculo computacional se va haciendo lento conforme el sistema crece en número de partículas, y se simplifica tomando solo la parte electrónica del hamiltoniano, de acuerdo con la aproximación de Born-Oppenheimer y el método de campo autoconsistente. De manera que quedó abierta la posibilidad de resolver reactividades químicas y apareció la química cuántica, la cual se ha apoyado en los postulados de la mecánica cuántica:

Postulado 1. El estado del sistema es completamente descrito por una función $\psi(r_1, r_2, \dots, t)$, conocida como función de onda de un sistema.

Postulado 2. Los observables son representados por operadores hermitianos que satisfacen las relaciones de conmutación.

$$\left[q, p_{q'}\right] = i\hbar\delta_{qq'} \qquad [q, q'] = 0 \qquad \left[p_q, p_{q'}\right] = 0$$

donde las q denotan las coordenadas de posición y las p_q las de momento lineal.

Postulado 3. Cuando un sistema queda descrito por una función de onda ψ el valor medio de la observable Ω en una serie de mediciones es igual al valor esperado del operador correspondiente

$$\langle\Omega\rangle = \frac{\int \psi \cdot \Omega\psi d\tau}{\int \psi \cdot \psi d\tau}$$

donde la función de onda es una combinación lineal de las eigenfunciones de Ω

$$\psi = \sum C_n\psi_n \qquad N\acute{o}tese\ que \quad \Omega\psi_n = \omega_n\psi_n$$

Postulado 4. La probabilidad de que una partícula sea encontrada en el elemento de volumen τ en un punto r es proporcional a: $|\psi(r)|^2 d\tau$

Postulado 5. La función de onda $\psi\left(\bar{r}_1, \bar{r}_2 \dots t\right)$ evoluciona con el tiempo de acuerdo con:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi$$

La aproximación de Born-Oppenheimer

Los orbitales moleculares y las superficies de energía de las interacciones átomo-átomo, átomo-molécula y molécula-molécula se pueden encontrar resolviendo la ecuación de onda de Schrödinger independiente del tiempo para sistemas multielectrónicos. Dicha ecuación de Schrödinger en notación de Dirac está dada por

$$H|\psi\rangle = E|\psi\rangle \tag{5}$$

donde H es el operador hamiltoniano de un sistema molecular (constituido de N electrones y S núcleos). E es la energía del sistema y $0\,\psi$ es la función de onda del sistema. El hamiltoniano completo (en unidades atómicas) está descrito por la siguiente expresión:

$$H = -\sum_{\mu=1}^{N}\frac{1}{2}\nabla_{\mu}^{2} - \sum_{k=1}^{S}\frac{\nabla_{k}^{2}}{2M_{k}} - \sum_{\mu=1}^{N}\sum_{k=1}^{S}\frac{Z_{k}}{r_{k\mu}} + \sum_{\mu=1}^{N}\sum_{\mu>\nu}^{N}\frac{1}{r_{\mu\nu}} + \sum_{k>l}^{S}\frac{Z_{k}Z_{l}}{R_{kl}} \tag{6}$$

En esta expresión, M_k es el cociente entre la masa del núcleo y la masa del electrón, y Z_k es el número atómico del núcleo *k.* Los operadores laplacianos ∇_{μ}^2 y ∇_{k}^2 indican diferenciación con respecto a las coordenadas del μ-ésimo electrón y del *k*-ésimo núcleo, respectivamente. La distancia entre él μ-ésimo electrón y el *k*-ésimo núcleo es $r_{\mu k}$; R_{kl}, representa la distancia entre el *k*-ésimo núcleo y el *l*-ésimo núcleo y $r_{\mu\nu}$ es la distancia entre μ-ésimo electrón y él ν-ésimo electrón. El primer término del hamiltoniano es el operador de energía cinética de los electrones, el segundo término representa el operador de energía cinética de los núcleos, el tercer término constituye la energía potencial de la atracción coulombiana entre los electrones y los núcleos, el cuarto es la energía potencial de repulsión entre electrones y el quinto término representa la energía potencial de repulsión entre núcleos. La ecuación (6) depende de las coordenadas de los núcleos y los electrones, por lo tanto, la solución de la ecuación de Schrödinger depende también de ambos conjuntos de coordenadas.

$$H\psi(r,R) = E\psi(r,R) \tag{7}$$

En donde *r* representa a las coordenadas electrónicas y *R* a las nucleares.

Como la masa del núcleo es ~1800 veces mayor que la masa del electrón al comparar el primero con el segundo término del hamiltoniano (6), se encuentra que el movimiento de los núcleos es mucho más lento que el de los electrones. Esto permite considerar el

movimiento de los electrones como si los núcleos permanecieran fijos en el espacio. Un ligero cambio en el movimiento de los núcleos deja prácticamente a los electrones en un estado estacionario.

A su vez, el movimiento de los núcleos puede describirse como el resultado de un campo promedio debido a los electrones actuando sobre estos. En términos de la función de onda, lo anterior significa que esta función puede escribirse como un producto de funciones

$$\psi(r,R) = \psi_R(r)\phi(R) \tag{8}$$

donde la función $\psi_R(r)$ es la función de onda electrónica para posiciones nucleares fijas, la cual depende solo de los estados cuánticos de los electrones; es decir, esta función depende paramétricamente de las coordenadas nucleares. La función $\phi(R)$ es la función de onda nuclear. Esta describe el movimiento rotacional y vibracional de los núcleos en un potencial proporcionado por los electrones.

Entonces, debido a la gran diferencia entre las masas de los electrones y de los núcleos, los electrones responden instantáneamente al desplazamiento del núcleo, y dado que los núcleos se mueven más lentamente, se consideran fijos, mientras que los electrones en una molécula se mueven en el campo de estos núcleos fijos. Por ejemplo, en el caso de una molécula diatómica la ecuación de Schrödinger se resuelve tomando los núcleos estáticos, y considerando solo el movimiento de sus nubes electrónicas. Soluciones a la ecuación de Schrödinger para diferentes arreglos geométricos permiten construir curvas de energía potencial, y aun para moléculas poliatómicas se construyen superficies de energía potencial (PES, por sus siglas en inglés). Este es el primer paso de la aproximación de Born-Oppenheimer, confiable en el estado base del sistema.

La ecuación de Schrödinger describe el movimiento de los electrones, considerando que los núcleos permanecen fijos viene dada por

$$\left\{\sum_{\mu=1}^{N}\left[-\frac{1}{2}\nabla_\mu^2 - \sum_{k=1}^{S}\frac{Z_k}{r_{\mu k}} + \sum_{\mu>\nu}^{N}\frac{1}{r_{\mu\nu}}\right] + \sum_{k>\ell}^{S}\frac{Z_k Z_\ell}{r_{k\ell}}\right\}\psi_R(r) = \mathcal{E}\psi_R(r) \tag{9}$$

A esta expresión se le conoce como la ecuación de Schrödinger electrónica.

Para separar el movimiento nuclear, la expresión para la función de onda dada por la ecuación (8) se sustituye en la ecuación de Schrödinger del sistema molecular, esto es

$$\left\{\sum_{\mu=1}^{N}\left[-\frac{1}{2}\nabla_{\mu}^{2}-\sum_{k=1}^{S}\frac{Z_k}{r_{\mu k}}+\sum_{\mu>\nu}^{N}\frac{1}{r_{\mu\nu}}\right]-\sum_{k=1}^{S}\frac{1}{2M_k}\nabla_{k}^{2}+\sum_{k>\ell}^{S}\frac{Z_k Z_\ell}{r_{k\ell}}\right\}\psi_R(r)\phi(R)=E\psi_R(r)\phi(R) \quad (10)$$

Para desarrollar la ecuación (10) se necesitan las siguientes derivaciones:

$$\nabla_k^2[\psi_R(r)\phi(R)]=\nabla_k\cdot\left[\nabla_k\left(\psi_R(r)\phi(R)\right)\right]$$

$$=\nabla_k\cdot\left[\left(\nabla_k\psi_R(r)\right)\phi(R)+\psi_R(r)\nabla_k\phi(R)\right]$$

$$\therefore \quad \nabla_k^2[\psi_R(r)\phi(R)]=\left(\nabla_k^2\psi_R(r)\right)\phi(R)+2\nabla_k\psi_R(r)\cdot\nabla_k\phi(R)+\psi_R(r)\nabla_k^2\phi(R) \quad (11)$$

$$\nabla_\mu^2[\psi_R(r)\phi(R)]=\nabla_\mu\cdot\left[\nabla_\mu\left(\psi_R(r)\phi(R)\right)\right]$$

$$=\nabla_\mu\cdot\left[\left(\nabla_\mu\psi_R(r)\right)\phi(R)\right]$$

$$=\phi(R)\nabla_\mu^2\psi_R(r) \quad (12)$$

Sustituyendo (11) y (12) en (10) se obtiene

$$-\sum_{k=1}^{S}\frac{1}{M_k}\left[\nabla_k\psi_R(r)\cdot\nabla_k\phi(R)+\frac{1}{2}\left(\nabla_k^2\psi_R(r)\right)\phi(R)\right]-\sum_{k=1}^{S}\frac{1}{2M_k}\psi_R(r)\nabla_k^2\phi(R)$$
$$-\sum_{\mu=1}^{N}\left[\left(\frac{1}{2}\nabla_\mu^2+\sum_{k=1}^{S}\frac{Z_k}{r_{\mu k}}\right)\psi_R(r)\right]\phi(R)+\left(\sum_{k>\ell}^{S}\frac{Z_k Z_\ell}{r_{k\ell}}+\sum_{\mu>\nu}^{N}\frac{1}{r_{\mu\nu}}\right)\psi_R(r)\phi(R)$$

$$=\mathcal{E}\psi_R(r)\phi(R) \quad (13)$$

dentro de la aproximación adiabática o de Born-Oppenheimer se considera que la función electrónica varía muy lentamente con respecto a las coordenadas nucleares, entonces es posible despreciar los términos de la primera sumatoria; por lo tanto, la ecuación de Schrödinger de un sistema molecular queda dada por

$$-\sum_{k=1}^{S}\frac{1}{2M_k}\psi_R(r)\nabla_k^2\phi(R)-\sum_{\mu=1}^{N}\left[\left(\frac{1}{2}\nabla_\mu^2+\sum_{k=1}^{S}\frac{Z_k}{r_{\mu k}}\right)\psi_R(r)\right]\phi(R)$$
$$+\left(\sum_{k>\ell}^{S}\frac{Z_k Z_\ell}{r_{k\ell}}+\sum_{\mu>\nu}^{N}\frac{1}{r_{\mu\nu}}\right)\psi_R(r)\phi(R)=\mathcal{E}\psi_R(r)\phi(R) \quad (14)$$

donde $\mathcal{E}$ es una energía aproximada.

Comparando la ecuación (14) con la ecuación (9) se ve que

$$-\sum_{k=1}^{S}\frac{1}{2M_k}\psi_R(r)\nabla_k^2\phi(R)+\mathcal{E}\psi_R(r)\phi(R)=\mathcal{E}\psi_R(r)\phi(R) \tag{15}$$

Factorizando la parte electrónica se tiene lo siguiente

$$\left\{-\sum_{k=1}^{S}\frac{1}{2M_k}\nabla_k^2\phi(R)+\mathcal{E}\phi(R)-\mathcal{E}\phi(R)\right\}\psi_R(r)=0 \tag{16}$$

como la función electrónica $\psi_R(r)$ no es cero, entonces

$$-\sum_{k=1}^{S}\frac{1}{2M_k}\nabla_k^2\phi(R)+\mathcal{E}\phi(R)=\mathcal{E}\phi(R) \tag{17}$$

de donde

$$\left[-\sum_{k=1}^{S}\frac{1}{2M_k}\nabla_k^2+\mathcal{E}\right]\phi(R)=\mathcal{E}\phi(R) \tag{18}$$

Al comparar con la ecuación para el átomo de hidrógeno se puede concluir que la ecuación (18) es la ecuación de Schrödinger que describe el movimiento de los núcleos en un potencial $V=\mathcal{E}(r)$, proporcionado por los electrones. En resumen, dentro de la aproximación adiabática las soluciones de las ecuaciones (9) y (18) describen los movimientos electrónicos y nucleares en un sistema molecular, respectivamente.

Bajo esta aproximación se despreció el operador de energía cinética de los núcleos, entonces la energía potencial de la repulsión de los núcleos se considera como una constante. Los términos que sobreviven a esta simplificación constituyen el hamiltoniano electrónico que describe el movimiento de *N* electrones en el campo producido por *S* núcleos fijos (Atkins y Friedman, 1996). Por lo tanto, esta ecuación corresponde únicamente a la contribución electrónica (hamiltoniano electrónico), y para considerar la energía total del sistema es necesario adicionar la energía de los núcleos:

$$H=-\sum_{\mu=1}^{N}\frac{1}{2}\nabla_\mu^2-\sum_{\mu=1}^{N}\sum_{k=1}^{S}\frac{Z_k}{r_{\mu k}}+\sum_{\mu=1}^{N}\sum_{\mu>\nu}^{N}\frac{1}{r_{\mu\nu}}. \tag{19}$$

Entonces la ecuación de Schrödinger electrónica está dada por

$$H_{elect}\psi_{elect}=E_{elect}\psi_{elect}\ , \tag{20}$$

donde ψ_{elect} es la función de onda electrónica

$$\psi_{elect} = \psi_{elect}\left(r_\mu, R_k\right), \tag{21}$$

la cual contiene información física acerca de los electrones de la molécula. Esta función depende explícitamente de las coordenadas electrónicas r_μ e implícitamente de las coordenadas R_k de los núcleos, de manera similar que la energía electrónica

$$E_{elect} = E_{elect}\left(R_k\right). \tag{22}$$

La energía total en la aproximación de núcleos fijos viene dada por

$$E_{total} = E_{elect} + \sum_{k=1}^{S}\sum_{l>k}^{S}\frac{Z_k Z_l}{R_{kl}}, \tag{23}$$

donde la repulsión nuclear es constante. Entonces, la ecuación (23) conforma el problema electrónico de una molécula.

Para tratar el movimiento de los núcleos, que es el segundo paso de la aproximación Born-Oppenheimer, se procede de la misma forma que en el problema electrónico: reemplazamos las coordenadas electrónicas por sus valores del promedio obtenido sobre la función de onda electrónica. Lo anterior genera un hamiltoniano nuclear para el movimiento de los núcleos dentro del campo promedio de los electrones.

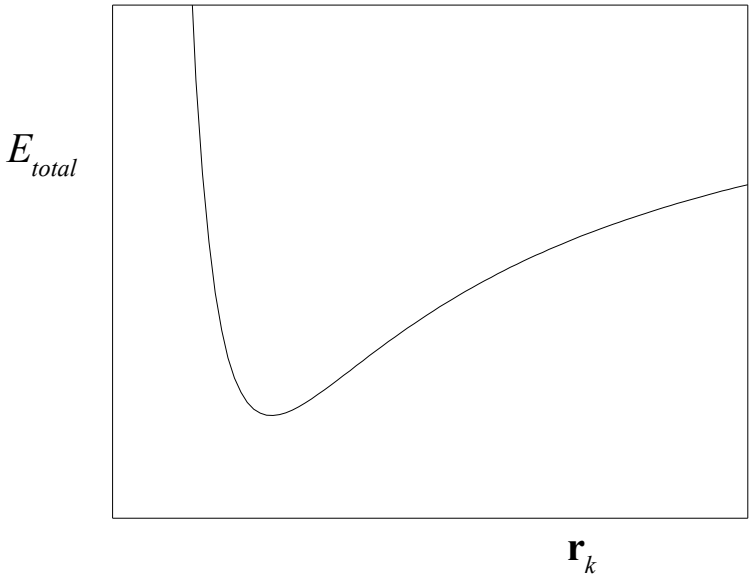
$$H_{nuclear} = -\sum_{k=1}^{S}\frac{\nabla_k^2}{2M_k} + \left\langle \; -\sum_{\mu=1}^{N}\frac{1}{2}\nabla_\mu^2 - \sum_{\mu=1}^{N}\sum_{k=1}^{S}\frac{Z_k}{r_{\mu k}} + \sum_{\mu=1}^{N}\sum_{\nu>1}^{N}\frac{1}{r_{\mu\nu}} \; \right\rangle + \sum_{k=1}^{S}\sum_{l>k}^{S}\frac{Z_k Z_l}{R_{kl}}$$

$$= \sum_{k=1}^{S}\frac{\nabla_k^2}{2M_k} + E_{elect}(R_k) + \sum_{k=1}^{S}\sum_{l>k}^{S}\frac{Z_k Z_l}{R_{kl}}$$

de donde $$H_{nuclear} = -\sum_{k=1}^{S}\frac{\nabla_k^2}{2M_k} + E_{total}(R_k) \tag{24}$$

La $E_{\text{total}}(R_k)$ proporciona una energía potencial para el movimiento de los núcleos. En general, esta función constituye una superficie de energía potencial, como se muestra esquemáticamente en la figura 1.

Figura 1. *Ilustración de una superficie de potencial*



Fuente: elaboración propia.

Las soluciones que proporciona la ecuación de Schrödinger nuclear:

$$H_{nuclear}\psi_{nuclear} = \xi\psi_{nuclear}, \tag{25}$$

describen la vibración, rotación y traslación de una molécula, donde la función de onda es

$$\psi_{nuclear} = \psi_{nuclear}(R_k), \tag{26}$$

y ξ es la aproximación de Born-Oppenheimer a la energía total de la ecuación (5), e incluye la energía electrónica, vibracional, rotacional y traslacional del sistema en consideración. La aproximación correspondiente a la función de onda de la ecuación (5) es

$$\psi(r_\mu, R_k) = \psi_{elect}(r_\mu, R_k)\psi_{nuclear}(R_k). \tag{27}$$

Debido a que en el primer paso de la aproximación de Born-Oppenheimer se considera que los núcleos *están fijos* (movimiento lento respecto a los electrones) y que la contribución por repulsión de los núcleos es *constante*, en lo que sigue se considerará solo la parte electrónica de la ecuación (23). En estas expresiones tanto el hamiltoniano electrónico como la función de onda electrónica se presentan sin el subíndice *elect*, esto por el uso de la aproximación de

Born-Oppenheimer; por lo tanto, la solución de la ecuación de Schrödinger electrónica está determinada por:

$$H\psi(\bar{r},\bar{R}) = E(\bar{R})\psi(\bar{r},\bar{R}), \tag{28}$$

para un conjunto fijo de núcleos localizados en $\bar{R}$. Así, el hamiltoniano está dado por:

$$H = -\sum_{\mu=1}^{N}\frac{1}{2}\nabla_{\mu}^{2} - \sum_{\mu=1}^{N}\sum_{k=1}^{S}\frac{Z_k}{r_{\mu k}} + \sum_{\mu=1}^{N}\sum_{\nu>1}^{N}\frac{1}{r_{\mu\nu}} + \text{cte.} \tag{29}$$

En cálculos de estructura molecular es conveniente no incluir la energía de repulsión núcleo-núcleo en H, y sumarla únicamente al final del cálculo como un término constante para cada configuración de los núcleos. Existen básicamente dos metodologías principales para resolver esta ecuación:

La primera metodología está relacionada con los cálculos *ab initio*, donde se escoge un modelo para la función de onda electrónica y se resuelve la ecuación (28), utilizando como entrada solo los valores de las constantes fundamentales y los números atómicos de los núcleos. La precisión está determinada fundamentalmente por el modelo seleccionado para dicha función de onda. Es importante mencionar que para moléculas mayores los cálculos *ab initio* son computacionalmente muy largos, dando lugar al desarrollo de los métodos conocidos como semiempíricos o aproximados para tratar una variedad más amplia de especies moleculares. El uso de pseudopotenciales permite mantener el estudio *ab initio* por lo menos de los electrones de valencia, esto es, aquellos que toman parte en la formación de las ligaduras del sistema molecular o supermolecular.

La segunda metodología tiene que ver precisamente con los métodos semiempíricos, en estos se hace uso de una forma simplificada del hamiltoniano, así como de parámetros ajustables con valores obtenidos de datos experimentales.

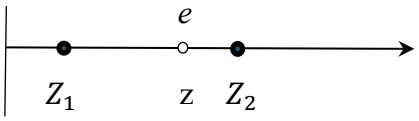
En ambos casos es una tarea extraordinaria calcular las energías llamadas químicamente precisas, es decir, energías con un error de 0.01 eV (~0.23 kcal/mol) de los valores exactos. Dado el gran número de integrales que se tienen que resolver con métodos *ab initio*, el método semiempírico manipula para despreciar o aproximar muchas de estas. Por ejemplo, la integral de traslape S la pone igual a la matriz identidad I. Esto implica, a saber, que la ecuación de Roothaan se use como transformación lineal fácil de manejar algebraicamente.

Como un ejemplo sencillo, consideremos el caso unidimensional de un ion molecular de hidrógeno con movimiento confinado al eje z, el hamiltoniano del sistema es

$$H = -\frac{\hbar^2}{2m_c}\frac{\partial^2}{\partial z^2} - \sum_{i=1,2}\frac{\hbar^2}{2m_i}\frac{\partial^2}{\partial Z_j^2} + V(z, Z_1, Z_2)$$

donde z es la localización del electrón y Z_i $(i = 1,2)$ la de los dos núcleos (figura 2).

Figura 2. *Localización del electrón entre dos núcleos en el eje z*



Fuente: elaboración propia.

El hamiltoniano es igual a la energía cinética del electrón más la energía cinética nuclear más la energía potencial del sistema.

$$H = T_c + T_N + V$$

La ecuación de Schrödinger es

$$H\Psi(z, Z_1, Z_2) = E\Psi(z, Z_1, Z_2)$$

Proponemos la solución

$$\Psi(z, Z_1, Z_2) = \psi(z; Z_1, Z_2)\chi(Z_1, Z_2)$$

Donde ψ es la función de onda electrónica y χ es la función de onda nuclear. Sustituyendo esta solución en la ecuación de Schrödinger se tiene:

$$H\psi\chi = \chi T_c\psi + \psi T_N\chi + V\psi\chi + W = E\psi\chi$$

donde

$$W = -\sum_{i=1,2}\frac{\hbar^2}{2m_i}\left(2\frac{\partial\psi}{\partial Z_i}\frac{\partial\chi}{\partial Z_i} + \frac{\partial^2\psi}{\partial Z_i^2}\chi\right)$$

W es responsable de los llamados *efectos no-adiabáticos*, que pueden ser muy importantes cuando las interacciones entre estados electrónicos son significativas. Esta última cantidad es distinta de cero porque ψ depende de las coordenadas nucleares, por lo que $\frac{\partial\psi}{\partial Z_i}$ también es

distinta de cero. Sin embargo, debido a que las masas nucleares se producen en el denominador, suponemos que W es pequeña y puede despreciarse, entonces

$$\chi T_c\psi + \psi T_N\chi + V\psi\chi = E\psi\chi$$

o bien,

$$\psi T_N\chi + (T_c\psi + V\psi)\chi = E\psi\chi$$

separando variables se tienen las ecuaciones:

$$T_c\psi + V\psi = E_c(Z_1, Z_2)\psi$$

$$T_N\chi + E_c\chi = E\chi$$

Las soluciones son la función de onda electrónica ψ con eigenvalor $E_c(Z_1, Z_2)$ y un potencial V que depende de la localización de los dos núcleos; y la función de onda χ de los núcleos cuando la energía potencial nuclear, ahora representada por E_c, tiene la forma de la curva de energía potencial molecular.

En este ejemplo se ve con claridad la separación de las ecuaciones diferenciales de Schrödinger, una para la parte electrónica y la otra para la parte nuclear.

Teoría orbital molecular

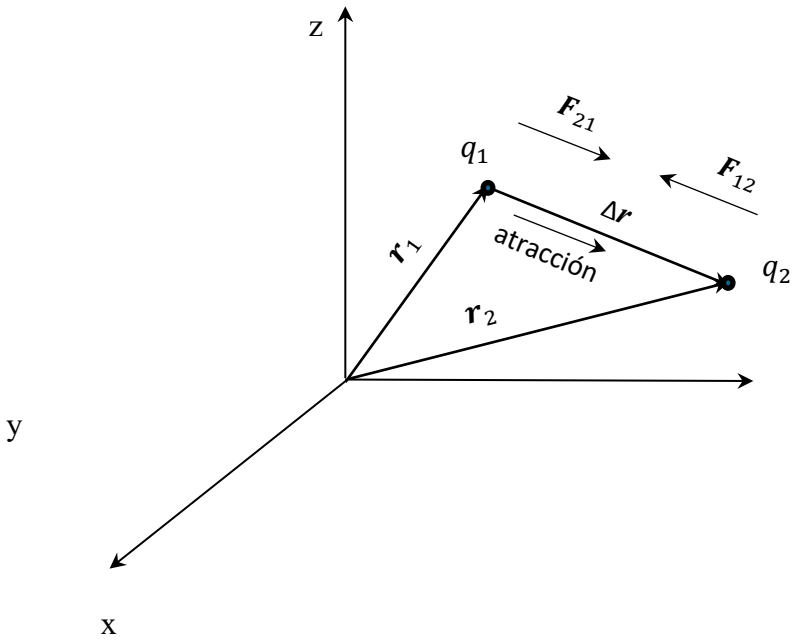
Un sistema molecular es un arreglo de núcleos y electrones que interactúan mutuamente mediante fuerzas electromagnéticas y cuyo comportamiento viene determinado por las leyes de la mecánica cuántica, por ejemplo, una molécula en un arreglo estructural entre moléculas.

La aproximación más cercana a la descripción teórica de sistemas moleculares es desde el punto de vista de la mecánica cuántica. De esta manera se pueden determinar propiedades electrónicas de átomos, moléculas o sistemas moleculares. El sistema más sencillo es el átomo de hidrógeno, el cual está constituido por dos partículas de cargas iguales y opuestas: un protón e^+ y un electrón e^-.

Sistema de dos partículas

Despreciando la interacción gravitacional sobre las dos partículas del átomo de hidrógeno, se tiene un sistema en el que las dos partículas interactúan mutuamente mediante la fuerza de Coulomb como se ve en la figura 3.

Figura 3. *Fuerzas de atracción entre cargas*



Fuente: elaboración propia.

La fuerza de Coulomb entre las dos partículas viene dada por

$$\boldsymbol{F}_{12} = \frac{kq_1q_2(\boldsymbol{r}_2 - \boldsymbol{r}_1)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^3} \qquad \boldsymbol{F}_{21} = \frac{kq_1q_2(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^3}$$

$$\boldsymbol{F}_{12} = -\boldsymbol{F}_{21}$$

$$\boldsymbol{r}_2 = \boldsymbol{r}_1 + \Delta\boldsymbol{r} \quad \Rightarrow \quad \Delta\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1$$

La fuerza total sobre el sistema de partículas viene dada por la suma vectorial de las fuerzas que actúan sobre cada una de las partículas

$$\boldsymbol{F} = \boldsymbol{F}_1 + \boldsymbol{F}_2$$

$$\boldsymbol{F}_1 = \boldsymbol{F}_{21} \qquad \boldsymbol{F}_2 = \boldsymbol{F}_{12}$$

Centro de masa: toda la fuerza y masa se concentran en el punto asociado al sistema de partículas que se mueve como si ahí se concentrara toda la masa y fuerza.

Por otro lado, sabemos que la posición del centro de masa de un sistema de partículas viene dada por

$$\boldsymbol{r}_{CM} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2}$$

donde m_1 y m_2 son respectivamente las masas de las partículas con vectores de posición $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$ con respecto a un origen de coordenadas. Derivando con respecto a t se encuentra la velocidad del centro de masa

$$\frac{d\boldsymbol{r}_{\boldsymbol{CM}}}{dt} = \frac{d}{dt}\left(\frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2}\right) \quad o\ bien \qquad \boldsymbol{v}_{CM} = \frac{m_1\boldsymbol{v}_1 + m_2\boldsymbol{v}_2}{m_1 + m_2}$$

Derivando nuevamente se encuentra la aceleración

$$\boldsymbol{a}_{CM} = \frac{m_1\boldsymbol{a}_1 + m_2\boldsymbol{a}_2}{m_1 + m_2}$$

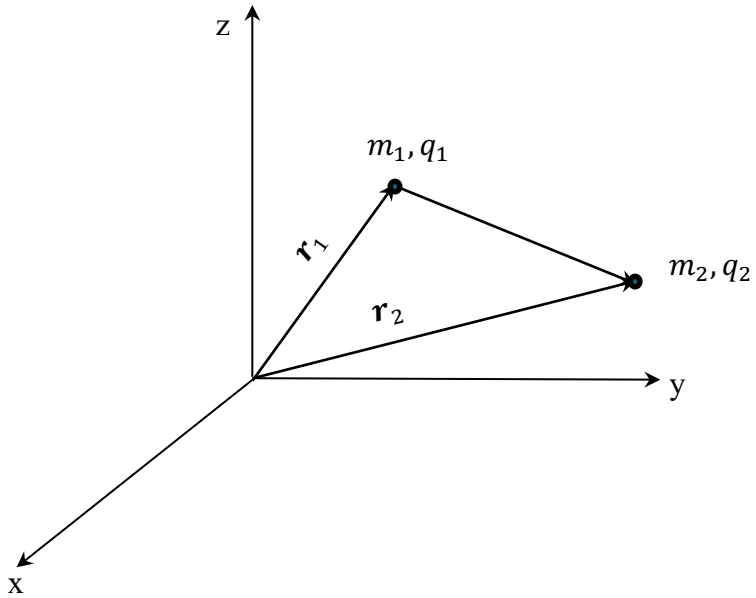
y como $m_1\boldsymbol{a}_1 = \boldsymbol{F}_1$ y $m_2\boldsymbol{a}_2 = \boldsymbol{F}_2$ la aceleración es

$$\boldsymbol{a}_{CM} = \frac{\boldsymbol{F}_1 + \boldsymbol{F}_2}{m_1 + m_2}$$

Nótese que si $\boldsymbol{F}_1 + \boldsymbol{F}_2 = 0$ entonces $\boldsymbol{a}_{CM} = 0$ y al integrar esta se obtiene $\boldsymbol{v}_{CM} = cte$.

Tomando un sistema de partículas con masas m_1 y m_2, respectivamente, como se ve en la figura 4.

Figura 4. *Sistema de dos partículas de masas m_1 y m_2, y cargas q_1 y q_2*



Fuente: elaboración propia.

la ecuación de Schrödinger viene dada por (las unidades de la ecuación son cgs):

$$\left[-\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\boldsymbol{r}_1 - \boldsymbol{r}_2)\right]\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = E\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$$

Donde V es la energía potencial del sistema de partículas debido a la fuerza de interacción (electrostática, fuerza de Coulomb).

Con el teorema de trabajo y energía se obtiene la fuerza para llevar la partícula 2 hasta la partícula 1

$$\boldsymbol{F}_{\boldsymbol{T_2}} = \boldsymbol{F}_{12} + \boldsymbol{F}_{ext}$$

$$W_{AB}(Total) = \int_C \boldsymbol{F}_T \cdot d\boldsymbol{r} = \int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} \boldsymbol{F}_T \cdot d\boldsymbol{r} = \frac{1}{2} m_2 (v_B^2 - v_A^2)$$

donde C es la trayectoria recorrida entre las partículas. La energía potencial es el mínimo trabajo externo para tener la configuración

$$\int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} (\boldsymbol{F}_{12} + \boldsymbol{F}_{ext}) \cdot d\boldsymbol{r} = \int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} \boldsymbol{F}_{12} \cdot d\boldsymbol{r} + \int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} \boldsymbol{F}_{ext} \cdot d\boldsymbol{r} = \frac{1}{2} m_2 (v_B^2 - v_A^2)$$

Despejando el trabajo externo se tiene que

$$\int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} \boldsymbol{F}_{ext} \cdot d\boldsymbol{r} = -\int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} \boldsymbol{F}_{12} \cdot d\boldsymbol{r} + \frac{1}{2} m_2 (v_B^2 - v_A^2)$$

si la velocidad es constante, entonces $v_B^2 - v_A^2 = 0$, y el trabajo es una fuerza conservativa que solo depende de los puntos finales, no de la trayectoria. Entonces, si $v = cte$

$$\int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} \boldsymbol{F}_{ext} \cdot d\boldsymbol{r} = -\int_{\boldsymbol{r}_A}^{\boldsymbol{r}_B} \boldsymbol{F}_{12} \cdot d\boldsymbol{r} = k \frac{q_1 q_2}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}$$
$$= V(\boldsymbol{r}_2 - \boldsymbol{r}_1)$$

como $k = 1$ en unidades cgs, por tanto

$$\left[-\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2 + \frac{q_1 q_2}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} \right] \psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = E\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$$

Se resuelve la ecuación de Schrödinger para encontrar las eigenfunciones ψ y los eigenvalores E. La ecuación de Schrödinger para el sistema de partículas viene dada por

$$\left[-\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2 + V_S \right] \psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = E\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$$

donde $V_S = \frac{q_1 q_2}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}$ es la energía potencial del sistema.

En nuestro caso, se desprecia la masa de las partículas, y la fuerza gravitacional; solo se consideran las fuerzas electrostáticas internas. Además, el centro de masa se mueve a$v = cte$.

Ahora bien, pasamos de una ecuación de seis variables a tres variables usando

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi(\boldsymbol{r}_{CM}, \boldsymbol{r}_1 - \boldsymbol{r}_2)$$

es decir, se pasa la ecuación de Schrödinger a términos de las variables relativas $(\boldsymbol{r}_1 - \boldsymbol{r}_2)$ y de las coordenadas del centro de masa (ver la figura 5).

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2 \qquad R = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2}$$

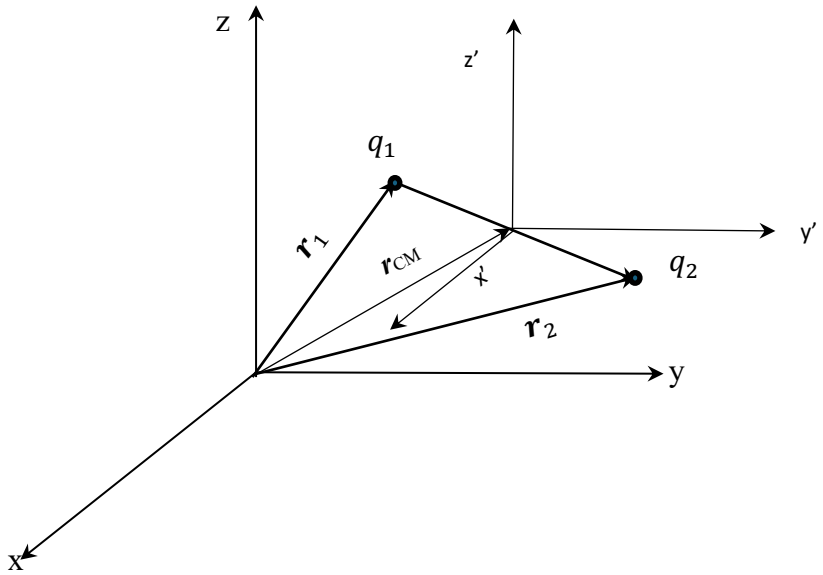
Explícitamente, estas variables son:

$$x = x_1 - x_2, \qquad y = y_1 - y_2, \qquad z = z_1 - z_2$$

$$X_R = \frac{m_1x_1 + m_2x_2}{m_1 + m_2}, \qquad Y_R = \frac{m_1y_1 + m_2y_2}{m_1 + m_2}, \qquad Z_R = \frac{m_1z_1 + m_2z_2}{m_1 + m_2}$$

con *R=CM*

Figura 5. *Sistema de referencia hacia el centro de masa*



Fuente: elaboración propia.

Considerando dos funciones $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ y $\Psi(\boldsymbol{r}, \boldsymbol{R})$ relacionadas por $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \Psi\big(\boldsymbol{r}(\boldsymbol{r}_1, \boldsymbol{r}_2), \boldsymbol{R}(\boldsymbol{r}_1, \boldsymbol{r}_2)\big)$. Y con el cambio de variable $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \mathbb{h}\big(\boldsymbol{r}(\boldsymbol{r}_1, \boldsymbol{r}_2), \boldsymbol{R}(\boldsymbol{r}_1, \boldsymbol{r}_2)\big)$ donde $(\boldsymbol{r}, \boldsymbol{R}) = \left(\boldsymbol{r}_1 - \boldsymbol{r}_2, \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2}\right)$ con $\mathbb{h}: R^6 \longrightarrow R^6$. Entonces, del cálculo de varias variables se tiene que $(\mathbb{D} \longrightarrow derivada)$

$$\mathbb{D}\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \mathbb{D}\Psi(\boldsymbol{r}, \boldsymbol{R})$$

El jacobiano de la transformación $\psi(x_1, y_1, z_1, x_2, y_2, z_2) \longrightarrow \Psi(x, y, z, X, Y, Z)$ está dado por

$$J(x_1, y_1, z_1, x_2, y_2, z_2) = J\left(\frac{x,y,z,X,Y,Z}{x_1,y_1,z_1,x_2,y_2,z_2}\right) = \begin{pmatrix} \frac{\partial x}{\partial x_1} & \frac{\partial x}{\partial y_1} & \frac{\partial x}{\partial z_1} & \frac{\partial x}{\partial x_2} & \frac{\partial x}{\partial y_2} & \frac{\partial x}{\partial z_2} \\ \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial y_1} & \frac{\partial y}{\partial z_1} & \frac{\partial y}{\partial x_2} & \frac{\partial y}{\partial y_2} & \frac{\partial y}{\partial z_2} \\ \frac{\partial z}{\partial x_1} & \frac{\partial z}{\partial y_1} & \frac{\partial z}{\partial z_1} & \frac{\partial z}{\partial x_2} & \frac{\partial z}{\partial y_2} & \frac{\partial z}{\partial z_2} \\ \frac{\partial X}{\partial x_1} & \frac{\partial X}{\partial y_1} & \frac{\partial X}{\partial z_1} & \frac{\partial X}{\partial x_2} & \frac{\partial X}{\partial y_2} & \frac{\partial X}{\partial z_2} \\ \frac{\partial Y}{\partial x_1} & \frac{\partial Y}{\partial y_1} & \frac{\partial Y}{\partial z_1} & \frac{\partial Y}{\partial x_2} & \frac{\partial Y}{\partial y_2} & \frac{\partial Y}{\partial z_2} \\ \frac{\partial Z}{\partial x_1} & \frac{\partial Z}{\partial y_1} & \frac{\partial Z}{\partial z_1} & \frac{\partial Z}{\partial x_2} & \frac{\partial Z}{\partial y_2} & \frac{\partial Z}{\partial z_2} \end{pmatrix}$$

Evaluando las derivadas parciales y definiendo $M = m_1 + m_2$ se tiene la siguiente matriz

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ \frac{m_1}{M} & 0 & 0 & \frac{m_2}{M} & 0 & 0 \\ 0 & \frac{m_1}{M} & 0 & 0 & \frac{m_2}{M} & 0 \\ 0 & 0 & \frac{m_1}{M} & 0 & 0 & \frac{m_2}{M} \end{pmatrix}$$

El producto del gradiente de la función de onda por el jacobiano nos da:

$$\nabla\Psi J(x_1, y_1, z_1, x_2, y_2, z_2)$$

$$= \left(\frac{\partial\Psi}{\partial x}, \frac{\partial\Psi}{\partial y}, \frac{\partial\Psi}{\partial z}, \frac{\partial\Psi}{\partial X}, \frac{\partial\Psi}{\partial Y}, \frac{\partial\Psi}{\partial Z}\right) \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ \frac{m_1}{M} & 0 & 0 & \frac{m_2}{M} & 0 & 0 \\ 0 & \frac{m_1}{M} & 0 & 0 & \frac{m_2}{M} & 0 \\ 0 & 0 & \frac{m_1}{M} & 0 & 0 & \frac{m_2}{M} \end{pmatrix}$$

$$= \left(\frac{\partial\Psi}{\partial x} + \frac{m_1}{M}\frac{\partial\Psi}{\partial X}, \frac{\partial\Psi}{\partial y} + \frac{m_1}{M}\frac{\partial\Psi}{\partial Y}, \frac{\partial\Psi}{\partial z} + \frac{m_1}{M}\frac{\partial\Psi}{\partial Z}, -\frac{\partial\Psi}{\partial x} + \frac{m_2}{M}\frac{\partial\Psi}{\partial X}, -\frac{\partial\Psi}{\partial y} + \frac{m_2}{M}\frac{\partial\Psi}{\partial Y}, -\frac{\partial\Psi}{\partial z} + \frac{m_2}{M}\frac{\partial\Psi}{\partial Z}\right)$$

Ahora bien

$$\frac{\partial\psi}{\partial x_1} = \frac{\partial\Psi}{\partial x} + \frac{m_1}{M}\frac{\partial\Psi}{\partial X_R}, \cdots \implies \nabla_{r_1}\psi = \nabla_{\boldsymbol{r}}\Psi + \frac{m_1}{M}\nabla_R\Psi$$

$$\frac{\partial\psi}{\partial x_2} = -\frac{\partial\Psi}{\partial x} + \frac{m_2}{M}\frac{\partial\Psi}{\partial X_R},\cdots \quad \Longrightarrow \qquad \nabla_{\boldsymbol{r}_2}\psi = -\nabla_{\boldsymbol{r}}\Psi + \frac{m_2}{M}\nabla_R\Psi$$

$$\nabla_{r_1}^2\psi = \nabla_{\boldsymbol{r}_1}\cdot\left[\nabla_{\boldsymbol{r}_1}\psi\right]$$

$$\nabla_{r_1}^2\psi = \left(\nabla_{\boldsymbol{r}}\Psi + \frac{m_1}{M}\nabla_R\right)\cdot\left[\nabla_{\boldsymbol{r}}\Psi + \frac{m_1}{M}\nabla_R\Psi\right]$$

$$= \nabla_r^2\Psi + \frac{2m_1}{M}\nabla_{\boldsymbol{r}}\cdot(\nabla_{\boldsymbol{r}}\Psi) + \left(\frac{m_1}{M}\right)^2\nabla_R^2\Psi$$

En forma análoga

$$\nabla_{r_2}^2\psi = \nabla_r^2\Psi - \frac{2m_2}{M}\nabla_{\boldsymbol{r}}\cdot(\nabla_{\boldsymbol{r}}\Psi) + \left(\frac{m_2}{M}\right)^2\nabla_R^2\Psi$$

Sustituyendo en la ecuación de Schrödinger, se tiene

$$-\frac{\hbar^2}{2m_1}\nabla_{r_1}^2\psi - \frac{\hbar^2}{2m_2}\nabla_{r_2}^2\psi = -\frac{\hbar^2}{2}\left[\frac{\nabla_{r_1}^2\psi}{m_1} + \frac{\nabla_{r_2}^2\psi}{m_2}\right]$$

$$= -\frac{\hbar^2}{2}\left[\frac{1}{m_1}\nabla_r^2\Psi + \frac{2}{M}\nabla_{\boldsymbol{r}}\cdot(\nabla_{\boldsymbol{r}}\Psi) + \frac{m_1}{M^2}\nabla_R^2\Psi + \frac{1}{m_2}\nabla_r^2\Psi - \frac{2}{M}\nabla_{\boldsymbol{r}}\cdot(\nabla_{\boldsymbol{r}}\Psi) + \frac{m_2}{M^2}\nabla_R^2\Psi\right]$$

$$= -\frac{\hbar^2}{2}\left[\left(\frac{1}{m_1} + \frac{1}{m_2}\right)\nabla_r^2\Psi + \left(\frac{m_1}{M^2} + \frac{m_2}{M^2}\right)\nabla_R^2\Psi\right]$$

$$= -\frac{\hbar^2}{2}\left[\left(\frac{1}{m_1} + \frac{1}{m_2}\right)\nabla_r^2\Psi + \left(\frac{1}{m_1 + m_2}\right)\nabla_R^2\Psi\right]$$

$$= -\frac{\hbar^2}{2}\left[\frac{1}{\mu}\nabla_r^2\Psi + \left(\frac{1}{m_1 + m_2}\right)\nabla_R^2\Psi\right]$$

Dado que $\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \Longrightarrow \mu = \frac{m_1 m_2}{m_1 + m_2}$. Reacomodando y arreglando se encuentra que

$$-\frac{\hbar^2}{2m_1}\nabla_{r_1}^2\psi - \frac{\hbar^2}{2m_2}\nabla_{r_2}^2\psi = -\frac{\hbar^2}{2}\left[\frac{1}{\mu}\nabla_r^2\Psi + \left(\frac{1}{m_1 + m_2}\right)\nabla_R^2\Psi\right]$$

Dado que $-\frac{\hbar^2}{2m_1}\nabla_{r_1}^2\psi - \frac{\hbar^2}{2m_2}\nabla_{r_2}^2\psi = -\frac{\hbar^2}{2M}\nabla_R^2\Psi - \frac{\hbar^2}{2\mu}\nabla_r^2\Psi$ se tiene que:

$$-\frac{\hbar^2}{2\mu}\nabla_r^2\Psi - \frac{\hbar^2}{2M}\nabla_R^2\Psi + \mathrm{V}_S\Psi = \mathrm{E}\Psi$$

donde $\mathrm{V}_S = \frac{q_1 q_2}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}$, $\qquad y \qquad r = |\boldsymbol{r}_2 - \boldsymbol{r}_1|$,

Separando variables

$$\Psi(\boldsymbol{r}, \boldsymbol{R}) = \mathcal{U}(\boldsymbol{r})\psi(\boldsymbol{R})$$

$$\nabla_r^2\Psi(\boldsymbol{r}, \boldsymbol{R}) = \psi(\boldsymbol{R})\nabla_r^2\mathcal{U}(\boldsymbol{r})$$

$$\nabla_R^2\Psi(\boldsymbol{r}, \boldsymbol{R}) = \mathcal{U}(\boldsymbol{r})\nabla_R^2\psi(\boldsymbol{R})$$

Si sustituimos en la ecuación de Schrödinger se tiene

$$-\frac{\hbar^2}{2\mu}\psi(\boldsymbol{R})\nabla_r^2\mathcal{U}(\boldsymbol{r})-\frac{\hbar^2}{2M}\mathcal{U}(\boldsymbol{r})\nabla_R^2\psi(\boldsymbol{R})+\mathrm{V}(\boldsymbol{r})\mathcal{U}(\boldsymbol{r})\psi(\boldsymbol{R})=\mathrm{E}\mathcal{U}(\boldsymbol{r})\psi(\boldsymbol{R})$$

Al multiplicar ambos miembros por $\frac{1}{\mathcal{U}(\boldsymbol{r})\psi(\boldsymbol{R})}$ se tiene que

$$-\frac{\hbar^2}{2\mu}\frac{1}{\mathcal{U}(\boldsymbol{r})}\nabla_r^2\mathcal{U}(\boldsymbol{r})-\frac{\hbar^2}{2M}\frac{1}{\psi(\boldsymbol{R})}\nabla_R^2\psi(\boldsymbol{R})+\mathrm{V}(\boldsymbol{r})-E=0$$

La energía total se puede separar en una energía radial y una del centro de masa

$$E=E_r+E_{CM}\qquad donde\quad E_{CM}=E_R$$

entonces

$$-\frac{\hbar^2}{2\mu}\frac{1}{\mathcal{U}(\boldsymbol{r})}\nabla_r^2\mathcal{U}(\boldsymbol{r})+\mathrm{V}(\boldsymbol{r})-E_r-\frac{\hbar^2}{2M}\frac{1}{\psi(\boldsymbol{R})}\nabla_R^2\psi(\boldsymbol{R})-E_R=0$$

dado que son independientes, se tienen las ecuaciones de Schrödinger radial y la del centro de masa:

$$-\frac{\hbar^2}{2\mu}\frac{1}{\mathcal{U}(\boldsymbol{r})}\nabla_r^2\mathcal{U}(\boldsymbol{r})+\mathrm{V}(\boldsymbol{r})-E_r=0$$

$$-\frac{\hbar^2}{2M}\frac{1}{\psi(\boldsymbol{R})}\nabla_R^2\psi(\boldsymbol{R})-E_R=0$$

Mientras la ecuación de Schrödinger del centro de masa

$$-\frac{\hbar^2}{2M}\nabla_R^2\psi(\boldsymbol{R})=E_R\psi(\boldsymbol{R})$$

se comporta como una partícula libre, la ecuación de Schrödinger de la parte radial es

$$-\frac{\hbar^2}{2\mu}\nabla_r^2\mathcal{U}(\boldsymbol{r})+\mathcal{U}(\boldsymbol{r})\mathrm{V}(\boldsymbol{r})=E_r\mathcal{U}(\boldsymbol{r})$$

Si $q_1=-e\quad y\quad q_2=Ze$, y como la masa m_2 del protón es mucho mayor que la masa m_1 del electrón, y dado que $\mu=\frac{m_1m_2}{m_1+m_2},\ si\ m_2>m_1\implies\mu=m_1$. Por lo tanto,

$$-\frac{\hbar^2}{2m_1}\nabla_r^2\mathcal{U}(\boldsymbol{r})-\frac{Ze^2}{r}\mathcal{U}(\boldsymbol{r})=E_r\mathcal{U}(\boldsymbol{r})$$

La ecuación de Schrödinger de un electrón e^- en presencia de un núcleo de número atómico Z, o bien la ecuación de Schrödinger del átomo de hidrógeno.

Las coordenadas naturales para resolver esta ecuación son las coordenadas esféricas (figura 6) con las condiciones de frontera $r\longrightarrow\infty$ cuando $\mathcal{U}\longrightarrow 0$. Al transformar de coordenadas cartesianas a coordenadas esféricas se obtiene el laplaciano:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial x^2} \implies \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2 sen\theta}\frac{\partial^2}{\partial \varphi^2} + \frac{1}{r^2} cot\theta \frac{\partial}{\partial \theta}$$

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 sen^2\theta}\frac{\partial^2}{\partial \varphi^2} + \frac{1}{r^2 sen\theta}\frac{\partial}{\partial \theta}\left(sen\theta \frac{\partial}{\partial \theta}\right)$$

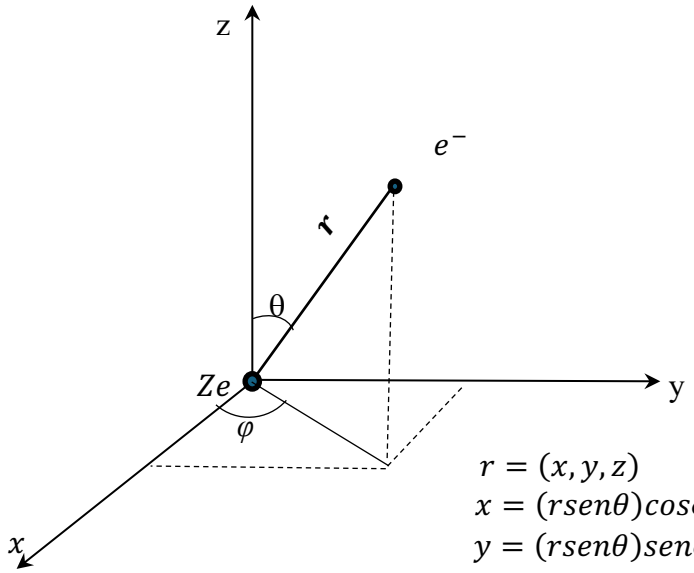
Aplicando este operador laplaciano a la ecuación de Schrödinger con $\mathcal{U}(\boldsymbol{r}) = \psi(r, \theta, \varphi)$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) + \frac{1}{r^2 sen^2\theta}\frac{\partial^2 \psi}{\partial \varphi^2} + \frac{1}{r^2 sen\theta}\frac{\partial}{\partial \theta}\left(sen\theta \frac{\partial \psi}{\partial \theta}\right) + \frac{2m_1}{\hbar^2}\left[E + \frac{Ze^2}{r}\right]\psi = 0$$

Se propone la solución

$$\psi(r, \theta, \varphi) = R(r)Y(\theta, \varphi)$$

Figura 6. *Transformación de coordenadas cartesianas a esféricas*



$$r = (x, y, z) \qquad coordenadas\ cartesianas$$
$$\left.\begin{aligned} x &= (rsen\theta)cos\varphi \\ y &= (rsen\theta)sen\varphi \\ z &= rcos\theta \end{aligned}\right\} \ coordenadas\ esféricas$$

Fuente: elaboración propia.

Sustituyendo el cambio de variable se tiene

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}[R(r)Y(\theta, \varphi)]\right) + \frac{1}{r^2 sen^2\theta}\frac{\partial^2}{\partial \varphi^2}[R(r)Y(\theta, \varphi)]$$
$$+ \frac{1}{r^2 sen\theta}\frac{\partial}{\partial \theta}\left(sen\theta \frac{\partial}{\partial \theta}[R(r)Y(\theta, \varphi)]\right) + \frac{2m_1}{\hbar^2}\left[E + \frac{Ze^2}{r}\right][R(r)Y(\theta, \varphi)] = 0$$

$$\frac{Y(\theta, \varphi)}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial R(r)}{\partial r}\right) + \frac{R(r)}{r^2 sen^2\theta}\frac{\partial^2 Y(\theta, \varphi)}{\partial \varphi^2}$$
$$+ \frac{R(r)}{r^2 sen\theta}\frac{\partial}{\partial \theta}\left(sen\theta \frac{\partial Y(\theta, \varphi)}{\partial \theta}\right) + \frac{2m_1}{\hbar^2}\left[E + \frac{Ze^2}{r}\right][R(r)Y(\theta, \varphi)] = 0$$

Multiplicamos esta expresión por el factor $\frac{r^2}{R(r)Y(\theta,\varphi)}$ y resulta

$$\frac{1}{R(r)}\frac{\partial}{\partial r}\left(r^2\frac{\partial R(r)}{\partial r}\right)+\frac{1}{Y(\theta,\varphi)}\frac{1}{sen^2\theta}\frac{\partial^2 Y(\theta,\varphi)}{\partial\varphi^2}$$

$$+\frac{1}{Y(\theta,\varphi)}\frac{1}{sen\theta}\frac{\partial}{\partial\theta}\left(sen\theta\frac{\partial Y(\theta,\varphi)}{\partial\theta}\right)+\frac{2m_1r^2}{\hbar^2}\left[E+\frac{Ze^2}{r}\right]=0$$

Dado que $\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right)=2r\frac{dR(r)}{dr}+r^2\frac{d^2}{dr^2}$, y al reacomodar se encuentra que:

$$\frac{r^2}{R(r)}\left(\frac{d^2R(r)}{dr^2}+\frac{2}{r}\frac{dR(r)}{dr}\right)+\frac{2m_1r^2}{\hbar^2}\left[E+\frac{Ze^2}{r}\right]$$

$$+\frac{1}{Y(\theta,\varphi)}\frac{1}{sen^2\theta}\frac{\partial^2 Y(\theta,\varphi)}{\partial\varphi^2}+\frac{1}{Y(\theta,\varphi)}\frac{1}{sen\theta}\frac{\partial}{\partial\theta}\left(sen\theta\frac{\partial Y(\theta,\varphi)}{\partial\theta}\right)=0$$

La parte de la ecuación que es función de la distancia se reconoce como $f(r)=\lambda$, y la que es función del ángulo como $q(\theta,\varphi)=-\lambda$, tal que al sumar estas dos funciones se anulan, esto es

$$f(r)+q(\theta,\varphi)=0$$

Con esto, la ecuación de Schrödinger se logra separar en dos ecuaciones: una radial y otra angular. Si multiplicamos por $Y(\theta,\varphi)$ la parte angular, se obtiene

$$\frac{\partial^2 Y(\theta,\varphi)}{\partial\theta^2}+cot\theta\frac{\partial Y(\theta,\varphi)}{\partial\theta}+\frac{1}{sen^2\theta}\frac{\partial^2 Y(\theta,\varphi)}{\partial\varphi^2}+\lambda Y(\theta,\varphi)=0$$

Se propone ahora

$$Y(\theta,\varphi)=T(\theta)$$

De manera que ahora, al sustituir, queda

$$F(\varphi)\frac{d^2T(\theta)}{d\theta^2}+F(\varphi)cot\theta\frac{dT(\theta)}{d\theta}+\frac{T(\theta)}{sen^2\theta}\frac{d^2F(\varphi)}{d\varphi^2}+\lambda T(\theta)F(\varphi)=0$$

Multiplicando por $\frac{sen^2\theta}{F(\varphi)T(\theta)}$ se puede separar variables como sigue

$$\frac{sen^2\theta}{T(\theta)}\frac{d^2T(\theta)}{d\theta^2}+\frac{sen^2\theta cot\theta}{T(\theta)}\frac{dT(\theta)}{d\theta}+sen^2\theta\,\lambda+\frac{1}{F(\varphi)}\frac{d^2F(\varphi)}{d\varphi^2}=0$$

La parte de la ecuación que es función de θ se reconoce como $g(\theta) = m$, y la que es función de φ como $h(\varphi) = -m$, tal que al sumar estas dos funciones se anulan, y con eso se satisface la ecuación diferencial, entonces

$$\frac{1}{F(\varphi)}\frac{d^2F(\varphi)}{d\varphi^2} = -m$$

$$\frac{sen^2\theta}{T(\theta)}\frac{d^2T(\theta)}{d\theta^2} + \frac{sen^2\theta cot\theta}{T(\theta)}\frac{dT(\theta)}{d\theta} + sen^2\theta\,\lambda = m$$

O bien

$$\frac{d^2F(\varphi)}{d\varphi^2} + mF(\varphi) = 0$$

$$\frac{d^2T(\theta)}{d\theta^2} + cot\theta\,\frac{dT(\theta)}{d\theta} + \left(\lambda - \frac{m}{sen^2\theta}\right)T(\theta) = 0$$

Si $m < 0$ la solución general es $F(\varphi) = Ae^{im\varphi}$

Nótese que

$$F(\varphi) = F(\varphi + 2\pi)$$

$$Ae^{im\varphi} = Ae^{im(\varphi+2\pi)}$$

$$Ae^{im\varphi} = Ae^{im\varphi}e^{2\pi mi} \quad \therefore \quad e^{2\pi mi} = 1$$

Por la relación de Euler $e^{i\theta} = cos\theta + isen\theta$ se tiene que

$$cos(2\pi m) + isen(2\pi m) = 1 \qquad \therefore \qquad m = 0, \pm1, \pm2, \pm3, \ \ldots$$

Si $m > 0$, y como m^2 siempre es positivo, entonces

$$\frac{d^2T(\theta)}{d\theta^2} + cot\theta\,\frac{dT(\theta)}{d\theta} + \left(\lambda - \frac{m^2}{sen^2\theta}\right)T(\theta) = 0$$

Si $x = cos\theta$ hacemos cambio de variable $T(\theta) = \mathrm{P}\big(x(\theta)\big)$ para poder resolver la ecuación diferencial

$$\frac{dT(\theta)}{d\theta} = \frac{d\mathrm{P}}{dx}\frac{dx}{d\theta} \quad \Longrightarrow \quad \frac{dT}{d\theta} = -sen\theta\,\frac{d\mathrm{P}}{dx}$$

$$cos^2\theta + sen^2\theta = 1 \implies sen^2\theta + x^2 = 1 \implies sen^2\theta = 1 - x^2$$

$$\therefore \qquad sen\theta = (1 - x^2)^{1/2}$$

$$\frac{dT}{d\theta} = -(1 - x^2)^{1/2}\frac{d\mathrm{P}}{dx}$$

$$\frac{d^2T}{d\theta^2} = \frac{d}{d\theta}\frac{dT}{d\theta} = -(1 - x^2)^{1/2}\frac{d}{dx}\left(-(1 - x^2)^{1/2}\frac{d\mathrm{P}}{dx}\right) \quad con$$

$$\frac{d}{d\theta} = -(1-x^2)^{1/2}\frac{d}{dx}$$

$$\frac{d^2T}{d\theta^2} = -\sqrt{1-x^2}\left(-\sqrt{1-x^2}\frac{d^2\mathrm{P}}{dx^2} + \frac{1}{2}\frac{2x}{\sqrt{1-x^2}}\frac{d\mathrm{P}}{dx}\right)$$

$$\frac{d^2T}{d\theta^2} = (1-x^2)\frac{d^2\mathrm{P}}{dx^2} - x\frac{d\mathrm{P}}{dx}$$

Finalmente, sustituyendo estos operadores en la ecuación diferencial se tiene:

$$(1-x^2)\frac{d^2\mathrm{P}}{dx^2} - x\frac{d\mathrm{P}}{dx} + \frac{x}{\sqrt{1-x^2}}\left(-(1-x^2)^{1/2}\frac{d\mathrm{P}}{dx}\right) + \left(\lambda - \frac{m^2}{sen^2\theta}\right)\mathrm{P} = 0$$

Es decir, se tiene la ecuación asociada de Legendre:

$$(1-x^2)\frac{d^2\mathrm{P}(x)}{dx^2} - 2x\frac{d\mathrm{P}(x)}{dx} + \left(\lambda - \frac{m^2}{sen^2\theta}\right)\mathrm{P}(x) = 0$$

para $-1 \le x \le 1 \quad y \quad m = 0, \pm1, \pm2, \pm3, \ \ldots \quad \therefore \quad si\ x = 1\ existe\ divergencia.$

$$\frac{d^2\mathrm{P}(x)}{dx^2} - \frac{2x}{(1-x^2)}\frac{d\mathrm{P}(x)}{dx} + \left(\frac{\lambda}{(1-x^2)} - \frac{m^2}{(1-x^2)^2}\right)\mathrm{P}(x) = 0$$

La ecuación es de la forma

$$y'' + P(x)y' + Q(x)y = 0$$

Lo usual es usar el método de potencias o método de Frobenius para resolver esta ecuación, sin embargo, si tiene puntos singulares (regulares o irregulares) se debe tener cuidado al proponer la solución en potencias (Arfken, 1985). Si P y Q permanecen finitas, son puntos ordinarios. Si P y Q divergen entonces son puntos singulares. Si tienen puntos singulares, entonces linealizamos en la forma

$$(x-x_0)P(x) \qquad y \qquad (x-x_0)^2Q(x)$$

En el límite cuando $x \to x_0$ y estas ecuaciones son finitas, entonces se tienen puntos singulares regulares. Por lo tanto, la ecuación asociada de Legendre tiene singularidad regular. Dado esto, se puede usar el método de potencias para resolverla. Este método se aplica en la forma

$$\mathrm{P}(x) = (1-x^2)^S G(x) \text{ donde } G(x) \neq 0$$

Para el átomo de hidrógeno lo que significa es que el electrón no esté en el núcleo. Y es bien sabido que la probabilidad de que esto suceda es casi cero.

$$\frac{d\mathrm{P}}{dx} = -2xS(1-x^2)^{S-1}G(x) + (1-x^2)^S\frac{dG}{dx}$$

$$\frac{d^2\mathrm{P}}{dx^2} = (1-x^2)^S\frac{d^2G}{dx^2} - 4Sx(1-x^2)^{S-1}\frac{dG}{dx} - 2S(1-x^2)^{S-1}G(x) + 4S(S-1)x^2(1-x^2)^{S-2}G(x)$$

Aplicando estos operadores y simplificando se obtiene

$$(1-x^2)^2\frac{d^2G}{dx^2} + (1-x^2)\left\{-4Sx\frac{dG}{dx} - 2SG(x) - 2x\frac{dG}{dx} + \lambda G(x)\right\} + \{4S(S-1)x^2G(x) + 4Sx^2G(x) - m^2G(x)\} = 0$$

Evaluando en $x = \pm 1$, solo quedan los términos que no tienen $(1-x^2)$, entonces

$$\{4S(S-1)G(x) + 4SG(x) - m^2G(x)\} = 0$$

Simplificando se obtiene:

$$4S^2 - m^2 = 0$$

$$S = \pm\frac{1}{2}|m|$$

Por lo tanto

$$\mathrm{P}(x) = (1-x^2)^{\pm\frac{1}{2}|m|}G(x)$$

Además, también se tiene que

$$+(1-x^2)\left\{(1-x^2)\frac{d^2G}{dx^2} - 4\left(\pm\frac{1}{2}|m|\right)x\frac{dG}{dx} - 2\left(\pm\frac{1}{2}|m|\right)G(x) - 2x\frac{dG}{dx} + \lambda G(x)\right\} + \left\{4\left(\pm\frac{1}{2}|m|\right)^2x^2G(x) - m^2G(x)\right\} = 0$$

$$(1-x^2)\left\{(1-x^2)\frac{d^2G}{dx^2} - 2(|m|+1)x\frac{dG}{dx} + (-|m|+\lambda)G(x)\right\} - (1-x^2)m^2G(x) = 0$$

$$\therefore \quad (1-x^2)\frac{d^2G}{dx^2} - 2(|m|+1)x\frac{dG}{dx} + (-|m| - m^2 + \lambda)G(x) = 0$$

Para resolver por series de potencias, proponemos:

$$G(x) = \sum_{n=0}^{\infty} a_n x^n$$

cuyas primera y segunda derivadas son:

$$\frac{dG(x)}{dx} = \sum_{n=0}^{\infty} n a_n x^{n-1}$$

$$\frac{d^2G(x)}{dx^2} = \sum_{n=0}^{\infty} n(n-1)a_n x^{n-2}$$

Sustituyendo en la ecuación diferencial se obtiene

$$\sum_{n=0}^{\infty} n(n-1)a_n x^{n-2} - \sum_{n=0}^{\infty} n(n-1)a_n x^n - 2(|m|+1)\sum_{n=0}^{\infty} n a_n x^n$$

$$+ \{\lambda - |m|(|m|+1)\}\sum_{n=0}^{\infty} a_n x^n = 0$$

Trasladamos la primera sumatoria a x^n haciendo el cambio de variable $n = N + 2$ y enseguida regresando a n=N, desde n=0, se adquiere

$$\sum_{n=0}^{\infty}\{(n+2)(n+1)a_{n+2} + [n(n-1) - 2(|m|+1)n + (\lambda - |m|(|m|+1))]a_n x^n\} = 0$$

Por independencia lineal se encuentra que

$$(n+2)(n+1)a_{n+2} + [n(n-1) - 2(|m|+1)n + (\lambda - |m|(|m|+1))]a_n = 0$$

Por lo tanto

$$a_{n+2} = \frac{(n+|m|)(n+|m|+1) - \lambda}{(n+2)(n+1)} a_n$$

$$\lim_{n\to\infty} \frac{a_{n+2}}{a_n} = \lim_{n\to\infty} \frac{(n+|m|)(n+|m|+1) - \lambda}{(n+2)(n+1)}$$

Dados $m \ y \ \lambda$ fijos

$$\frac{a_{n+2}}{a_n} = \frac{(n)n - \lambda}{(n)(n)} = 1$$

Pero si

$$\lim_{n\to\infty} \frac{a_{n+2}}{a_n} < 1$$

hay peligro de que diverja. Para reparar esto cortamos la serie. Esto se consigue haciendo

$$(n+|m|)(n+|m|+1) - \lambda = 0$$

$$\Longrightarrow \qquad \lambda = (n_r + |m|)(n_r + |m| + 1)$$

Por lo tanto λ es un entero positivo. Definiendo $\ell = n_r + |m|$ se tiene que $\lambda = \ell(\ell+1)$. Como ℓ es un entero positivo, entonces $|m| \leq \ell$. Por lo tanto,

$$(1-x^2)\frac{d^2\mathrm{P}}{dx^2} - 2(|m|+1)x\frac{d\mathrm{P}}{dx} + \big(\ell(\ell+1) - |m|(|m|+1)\big)\mathrm{P}(x) = 0$$

Comparable con la ecuación asociada de Legendre:

$$(1-x^2)\frac{d^2G}{dx^2}-2x\frac{dG}{dx}+(\ell(\ell+1)-m^2)G(x)=0$$

ℓ entero positivo y $\ell \geq |m|$ para m entero.

$$-\frac{\hbar^2}{2m_1}\nabla_r^2\psi-\frac{Ze^2}{r}\psi=E_r\psi$$

$$\psi(r,\theta,\varphi)=R(r)T(\theta)F(\varphi)$$

$$Y(\theta,\varphi)=T(\theta)F(\varphi)$$

$$T(\theta)=\mathrm{P}(x)$$

$$F(\varphi)=Be^{im\varphi} \qquad\qquad m=0,\pm1,\pm2,\pm3,\ \ldots$$

Las soluciones de la ecuación asociada de Legendre vienen dadas por

$$T(\theta)=A_\ell^{|m|}\mathrm{P}_\ell^{|m|}(x)$$

Donde $\mathrm{P}_\ell^{|m|}(x)$ es la fórmula de Rodrigues

$$\mathrm{P}_\ell^{|m|}(x)=\frac{1}{2^\ell\ell!}(1-x^2)^{|m|/2}\frac{d^{\ell+|m|}}{dx^{\ell+|m|}}(x^2+1)^\ell$$

$$A_\ell^{|m|}=\left[\frac{(2\ell+1)(\ell-|m|)!}{2(\ell+|m|)!}\right]^{1/2}$$

$$T(\theta)=\left[\frac{(2\ell+1)(\ell-|m|)!}{2(\ell+|m|)!}\right]^{1/2}\mathrm{P}_\ell^{|m|}(x)$$

Se normaliza pidiendo la condición:

$$\int \psi^*(r,\theta,\varphi)\,\psi(r,\theta,\varphi)dV=1$$

$$\int_0^\infty\int_0^{2\pi}\int_0^{\pi}R^*(r)R(r)T^*(\theta)T(\theta)F^*(\varphi)F(\varphi)r^2sen\theta dr d\theta d\varphi=1$$

$$\int_0^\infty R^2(r)r^2dr\int_0^{2\pi}F^2(\varphi)d\varphi\int_0^{\pi}T^2(\theta)sen\theta d\theta=1$$

Cada una de estas tres integrales es igual a 1 para que se satisfagan la igualdad y la normalización, particularmente, de acuerdo con un resultado anterior

$$\int_0^{2\pi}F^2(\varphi)d\varphi=\int_0^{2\pi}\left(B^*e^{-im\varphi}\right)\left(Be^{im\varphi}\right)d\varphi$$

$$=\int_0^{2\pi}B^*Bd\varphi$$

$$= B^2 \int_0^{2\pi} d\varphi$$

$$= B^2 2\,\pi = 1$$

$$\therefore \quad B = \frac{1}{\sqrt{2\pi}} \qquad es\ decir$$

$$F(\varphi) = \frac{1}{\sqrt{2\pi}} e^{im\varphi} \qquad para\ \ m = 0,1,2,\ \ .\ .\ .$$

Por lo tanto, se obtienen funciones armónicas esféricas:

$$Y_\ell^m(\theta,\varphi) = \left[\frac{2\ell+1}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}\right]^{1/2} \mathrm{P}_\ell^{|m|}(cos\theta)e^{im\varphi}$$

Nota: el término $e^{im\varphi}$ es la parte armónica y, además, conduce a un número complejo, todo lo demás es esférico. Recordando que ℓ es un entero positivo y que $\ell \geq |m|$ para m entero, analicemos caso por caso para valores más bajos de $\ell\ y\ m$. Definiendo $Y_\ell^m(\theta,\varphi) = Y_{\ell m}(\theta,\varphi)$ y evaluando $\mathrm{P}_\ell^{|m|}(x)$ con la fórmula de Rodrigues:

Caso $\ell = 0,\ \ m = 0$

$$Y_{00}(\theta,\varphi) = \frac{1}{\sqrt{4\pi}}$$

Caso $\ell = 1 \Longrightarrow \begin{cases} m = 1 \\ m = 0 \end{cases}$, para $m = 1$, la fórmula de Rodrigues lleva a

$$\mathrm{P}_1^1(cos\theta) = \frac{1}{2}(1 - cos^2\theta)^{1/2}\frac{d^2}{dx^2}(x^2+1)^1 = \frac{1}{2}(sen^2\theta)^{1/2}2 = sen\theta$$

$$Y_{11}(\theta,\varphi) = \sqrt{\frac{3}{8\pi}}sen\theta e^{i\varphi}$$

La solución es un número complejo, sin embargo $Y_{\ell m}(\theta,\varphi)$ es el producto de $e^{im\varphi}sen^{|m|}\theta$ y un polinomio de grado $(\ell - |m|)$ en $cos\theta$ y paridad $(-)^{\ell-m}$. En particular

$$Y_{\ell\ell} = (-)^\ell \left[\frac{2\ell+1}{4\pi}\frac{(2\ell)!}{2^{2\ell}(\ell!)^2}\right]^{1/2} sen^\ell\theta e^{i\ell\varphi}$$

$$\therefore \quad Y_{11}(\theta,\varphi) = -\sqrt{\frac{3}{8\pi}}sen\theta$$

Para poder obtener la parte real de los armónicos esféricos se recurre a las coordenadas esféricas, proponiendo como solución

$$\mathcal{U} = x^{\alpha} y^{\beta} z^{\gamma}$$

$$\mathcal{U} = (rsen\theta cos\varphi)^{\alpha}(rsen\theta sen\varphi)^{\beta}(rcos\theta)^{\gamma}$$

$$\mathcal{U} = r^{\alpha} r^{\beta} r^{\gamma}(sen\theta cos\varphi)^{\alpha}(sen\theta sen\varphi)^{\beta}(cos\theta)^{\gamma}$$

$$\mathcal{U} = r^{\alpha} r^{\beta} r^{\gamma} g(\theta, \varphi)$$

$$\mathcal{U} = r^{\alpha+\beta+\gamma} g(\theta, \varphi)$$

$$\alpha + \beta + \gamma = \ell$$

$$\mathcal{U}(r, \theta, \varphi) = r^{\ell} g(\theta, \varphi)$$

Este tipo de funciones satisface la ecuación de Laplace

$$\nabla^2 \mathcal{U} = 0$$

Expresando esta ecuación en términos de coordenadas esféricas se tiene lo siguiente

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2 sen^2\theta}\frac{\partial^2}{\partial \varphi^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2} cot\theta \frac{\partial}{\partial \theta}\right)\left(r^{\ell} g(\theta, \varphi)\right) = 0$$

Separando los términos radial y angular, el término radial se puede expresar como

$$\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right]\left(r^{\ell} g(\theta, \varphi)\right) = \ell(\ell - 1) r^{\ell-2}\, g(\theta, \varphi) + \frac{2\ell}{r} r^{\ell-1} g(\theta, \varphi)$$

$$= \ell(\ell + 1) r^{\ell-2} g(\theta, \varphi)$$

Mientras que el término angular se define como

$$\frac{1}{r^2}\left(\frac{1}{sen^2\theta}\frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial \theta^2} + cot\theta \frac{\partial}{\partial \theta}\right)\left(r^{\ell} g(\theta, \varphi)\right) = \frac{1}{r^2} r^{\ell-2} g(\theta, \varphi)\left(r^{\ell} g(\theta, \varphi)\right)$$

donde $\frac{1}{sen^2\theta}\frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial \theta^2} + cot\theta \frac{\partial}{\partial \theta} = \mathcal{L}$, entonces

$$\frac{1}{r^2}\mathcal{L}\left(r^{\ell} g(\theta, \varphi)\right) = r^{\ell-2}\mathcal{L} g(\theta, \varphi)$$

Sumando las contribuciones correspondientes, se tiene

$$\ell(\ell + 1) r^{\ell-2} g(\theta, \varphi) + r^{\ell-2}\mathcal{L} g(\theta, \varphi) = 0$$

Al simplificar y reacomodar se tiene que

$$\mathcal{L} g(\theta, \varphi) = -\ell(\ell + 1) g(\theta, \varphi)$$

Ahora bien, en la parte angular del átomo de hidrógeno se obtuvo

$$\frac{\partial^2 Y}{\partial \theta^2} + cot\theta \frac{\partial Y}{\partial \theta} + \frac{1}{sen^2\theta}\frac{\partial^2 Y}{\partial \varphi^2} + \lambda Y = 0$$

$$\mathcal{L} Y = -\lambda Y$$

$$\therefore \quad \lambda = \ell(\ell + 1)$$

Esto demuestra que se ha encontrado un método efectivo para determinar funciones que satisfacen la ecuación de Laplace para los armónicos esféricos. Esto es

$$\mathcal{U} = x^{\alpha} y^{\beta} z^{\gamma} = r^{\ell} g(\theta, \varphi)$$

$$\Longrightarrow g(\theta, \varphi) = \frac{x^{\alpha} y^{\beta} z^{\gamma}}{r^{\ell}}$$

es decir, esta ecuación se satisface siempre que $x^{\alpha} y^{\beta} z^{\gamma}$ satisfaga la ecuación de Laplace. Esto se consigue cuando $\ell = \alpha + \beta + \gamma$ de acuerdo con los desarrollos anteriores. Renombrando a la función $g(\theta, \varphi)$ como la función $Y_{\ell}(\theta, \varphi)$ se encuentra que

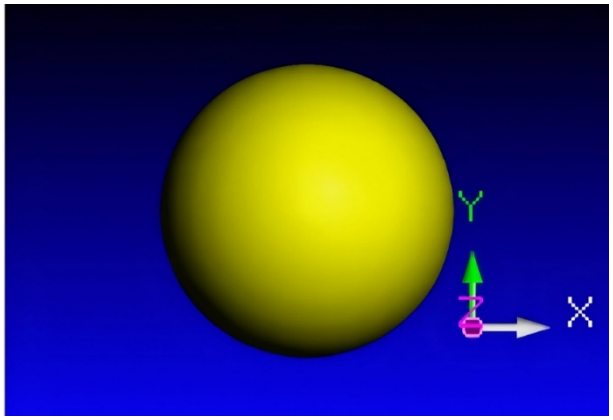
$$Y_{\ell}(\theta, \varphi) = A \frac{x^{\alpha} y^{\beta} z^{\gamma}}{r^{\ell}}$$

Para ℓ entero positivo (figura 7):

Si $\ell = 0$ se tiene que $\alpha + \beta + \gamma = 0$ para $\alpha = \beta = \gamma = 0$, entonces

$$Y_0(\theta, \varphi) = A$$

Figura 7. *Orbital s*



Fuente: elaboración propia.

Ahora bien, normalizando la función Y_0, esto es

$$\int_0^{2\pi} \int_0^{\pi} Y_{\ell}^{*} Y_{\ell}(\theta, \varphi) sen\theta d\theta d\varphi = 1$$

$$\int_0^{2\pi} \int_0^{\pi} A^2 sen\theta d\theta d\varphi = 1 \Longrightarrow A^2 2\pi \int_0^{\pi} sen\theta d\theta = 1$$

$$A^2 2\pi(2) = 1 \Longrightarrow A = \frac{1}{\sqrt{4\pi}}$$

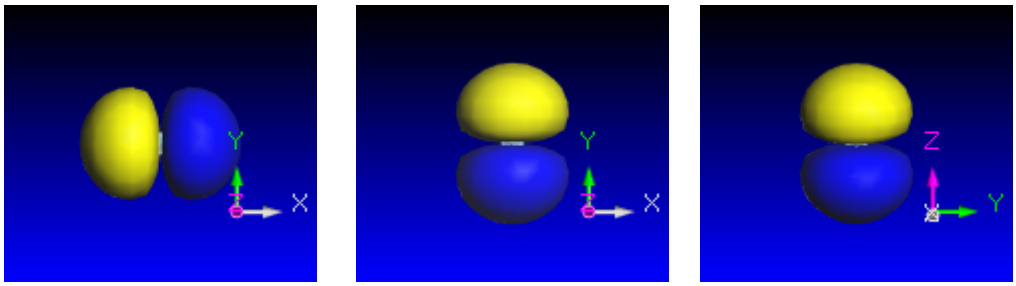
$$\therefore \qquad Y_0(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}$$

Si $\ell = 1$ se tienen tres tipos diferentes de funciones $x^{\alpha}y^{\beta}z^{\gamma}$ tales que $\alpha + \beta + \gamma = 1$, esto es $\left.\begin{matrix}\alpha = 1, \beta = 0, \gamma = 0 \\ \alpha = 0, \beta = 1, \gamma = 0 \\ \alpha = 0, \beta = 0, \gamma = 1\end{matrix}\right\}$ que cumplan $g(\theta, \varphi) = A\frac{x^{\alpha}y^{\beta}z^{\gamma}}{r^{\ell}}$, es decir (figura 8):

$$\begin{cases} g(\theta,\varphi) = A\frac{x}{r} = Asen\theta cos\varphi \rightarrow orbitales\ p_x \\ g(\theta,\varphi) = A\frac{y}{r} = Asen\theta sen\varphi \rightarrow orbitales\ p_y \\ g(\theta,\varphi) = A\frac{z}{r} = Acos\theta \quad \rightarrow orbitales\ p_z \end{cases}$$

En la normalización para $\ell = 1$ se encuentra que $A = \sqrt{\frac{3}{8\pi}}$.

Figura 8. *Orbitales p: a) p_x, b) p_y, c) p_z*



a *b* *c*

Fuente: elaboración propia.

Si $\ell = 2$, tomemos las siguientes posibilidades

$$\left.\begin{matrix}\alpha = 2, \beta = 0, \gamma = 0 \\ \alpha = 0, \beta = 2, \gamma = 0 \\ \alpha = 0, \beta = 0, \gamma = 2\end{matrix}\right\} \Longrightarrow \begin{cases} g(\theta,\varphi) = A\frac{x^2}{r} \\ g(\theta,\varphi) = A\frac{y^2}{r} \\ g(\theta,\varphi) = A\frac{z^2}{r} \end{cases}$$

Estas funciones no satisfacen la ecuación de Laplace, ya que al derivar es distinta de cero, por ejemplo:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)x^2 = 2$$

Ya que existen $2\ell + 1 = S$ armónicos esféricos, y dado que dos funciones linealmente independientes pueden formar una combinación lineal, entonces las ecuaciones que sí satisfacen la ecuación de Laplace son:

$$x^2 - y^2 \qquad y^2 - z^2 \qquad y \qquad x^2 - z^2$$

Ahora bien, entre estas elegimos arbitrariamente dos funciones linealmente independientes para construir una nueva combinación lineal:

$$x^2 - z^2 + y^2 - z^2 = x^2 + y^2 - 2z^2$$
$$= x^2 + y^2 + z^2 - 3z^2$$
$$= r^2 - 3z^2$$

a esta última expresión se le denomina usualmente d_{z^2}

Las cinco funciones linealmente independientes para $\ell = 2$, que satisfacen las ecuaciones de Laplace con sus correspondientes armónicos esféricos vienen dadas por (ver figura 9):

$$r^2 - 3z^2 \longrightarrow d_{z^2} \longrightarrow Y_{210} = \frac{1}{4}\left(\frac{5}{\pi}\right)^{1/2}(1 - 3cos^2\theta)$$

$$x^2 - y^2 \longrightarrow d_{x^2-y^2} \longrightarrow Y_{21}cos2\varphi = \frac{1}{2}\left(\frac{15}{\pi}\right)^{1/2} sen^2\theta cos2\varphi$$

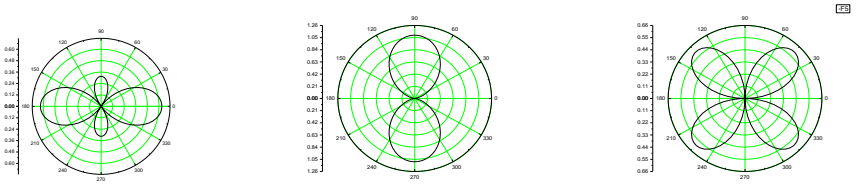
$$xy \longrightarrow d_{xy} \longrightarrow Y_{21}sen2\varphi = \frac{1}{2}\left(\frac{15}{\pi}\right)^{1/2} sen^2\theta sen2\varphi$$

$$xz \longrightarrow d_{xz} \longrightarrow Y_{21}cos\varphi = \frac{1}{4}\left(\frac{15}{\pi}\right)^{1/2} sen\theta cos\theta cos\varphi$$

$$yz \longrightarrow d_{yz} \longrightarrow Y_{21}sen\varphi = \frac{1}{4}\left(\frac{15}{\pi}\right)^{1/2} sen\theta cos\theta sen\varphi$$

Estos orbitales se representan polarmente en dos dimensiones de la siguiente forma (figura 9):

Figura 9. *Orbitales den forma polar*

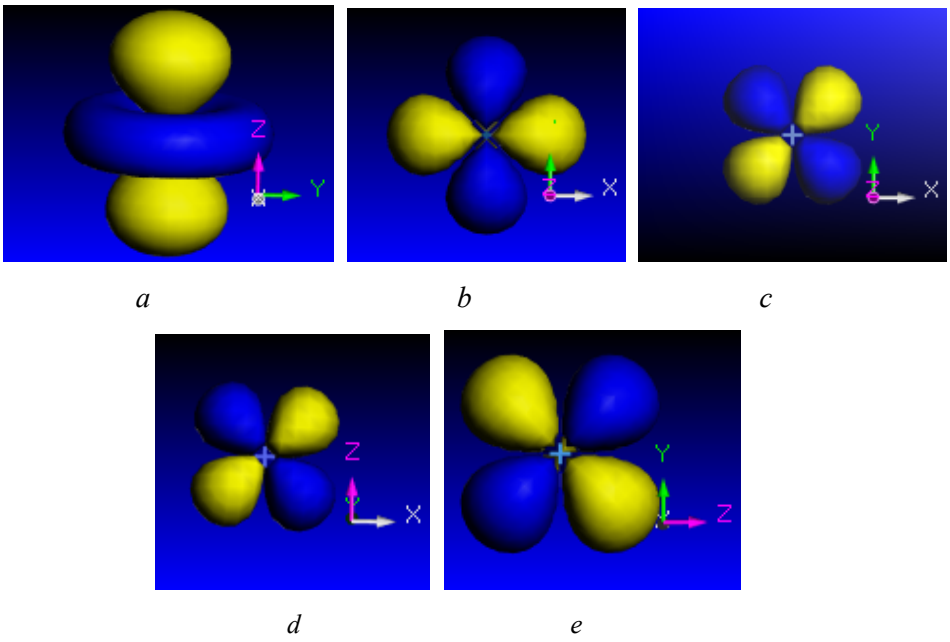


Y_{210} para d_{z^2} Y_{21} para $d_{x^2-y^2}$ y d_{xy} Y_{21} para d_{xz} y d_{yz}

Fuente: elaboración propia.

En tres dimensiones, los orbitales se representan en la figura 10 (las figuras tridimensionales de los orbitales han sido obtenidas usando el *software* BIOVIA Materials Studio [Dassault Systemes Biovia Corp., 2017):

Figura 10. *Orbitales d: a)* d_{z^2} *b)* $d_{x^2-y^2}$*, c)* d_{xy}*, d)* d_{xz}*, e)* d_{yz}



a *b* *c*

d *e*

Fuente: elaboración propia.

Cuando se utilizan los armónicos esféricos reales para describir la estructura electrónica de átomos y moléculas, a los valores $\ell = 0, 1, 2, 3, \ldots$ también se les identifica como orbitales $\ell = s, p, d, f, \ \ldots \ ,$ respectivamente. Estos se emplean para describir la estructura electrónica de los átomos. La ecuación radial está dada por

$$\frac{r^2}{R(r)}\left(\frac{d^2R(r)}{dr^2}+\frac{2}{r}\frac{dR(r)}{dr}\right)+\frac{2m_1r^2}{\hbar^2}\left[E+\frac{Ze^2}{r}\right]=\lambda$$

Multiplicando la expresión anterior por $\frac{R(r)}{r^2}$ se tiene

$$\frac{d^2R(r)}{dr^2}+\frac{2}{r}\frac{dR(r)}{dr}+\frac{2m_1}{\hbar^2}\left(E+\frac{Ze^2}{r}\right)R(r)=\lambda\frac{R(r)}{r^2}$$

Dado que $\lambda = \ell(\ell+1)$ donde $\ell > 0$ y $\ell \epsilon \mathbb{Z}^+$, reordenando se obtiene

$$\frac{d^2R(r)}{dr^2}+\frac{2}{r}\frac{dR(r)}{dr}+\frac{2m_1}{\hbar^2}\left(E+\frac{Ze^2}{r}-\frac{1}{2m_1}\frac{\ell(\ell+1)\hbar^2}{r^2}\right)R(r)=0$$

Cuando el electrón está capturado en el núcleo entonces se debe cumplir que

$$E < 0$$

Bajo esta consideración se tiene que

$$\frac{d^2R(r)}{dr^2}+\frac{2}{r}\frac{dR(r)}{dr}+\frac{2m_1}{\hbar^2}\left(-|E|+\frac{Ze^2}{r}-\frac{1}{2m_1}\frac{\ell(\ell+1)\hbar^2}{r^2}\right)R(r)=0$$

Al hacer el cambio de variable

$$\rho = \alpha r \qquad donde \qquad \alpha = \left(\frac{8m_1|E|}{\hbar^2}\right)^{1/2}$$

Este cambio de variable es para hacer adimensional la ecuación, esto es:

$$\frac{d^2\mathbb{R}}{d\rho^2} + \frac{2}{\rho}\frac{d\mathbb{R}}{d\rho} - \frac{\ell(\ell+1)}{\rho^2}\mathbb{R} - \frac{1}{4}\mathbb{R} + \frac{Ze^2}{\hbar}\left(\frac{m_1}{2|E|}\right)^{1/2}\frac{1}{\rho}\mathbb{R} = 0$$

Definiendo

$$\beta = \frac{Ze^2}{\hbar}\left(\frac{m_1}{2|E|}\right)^{1/2}$$

Al sustituir y reacomodar se obtiene la ecuación radial adimensional del átomo de hidrógeno:

$$\frac{d^2\mathbb{R}}{d\rho^2} + \frac{2}{\rho}\frac{d\mathbb{R}}{d\rho} - \frac{\ell(\ell+1)}{\rho^2}\mathbb{R} + \left(\frac{\beta}{\rho} - \frac{1}{4}\right)\mathbb{R} = 0$$

Esta ecuación tiene singularidad en $\rho = 0$, de modo que si $\rho \longrightarrow \infty$:

$$\frac{d^2\mathbb{R}}{d\rho^2} - \frac{1}{4}\mathbb{R} = 0$$

La ecuación característica es $m^2 - \frac{1}{4}m = 0$, cuyas soluciones $m = \pm\frac{1}{2}$ indican la solución a la ecuación diferencial es una combinación lineal de dos soluciones: $\mathbb{R} = Ae^{\frac{1}{2}\rho} + Be^{-\frac{1}{2}\rho}$. Cada una de estas dos soluciones son soluciones a la ecuación radial del átomo de hidrógeno; sin embargo, solo una de ellas satisface la condición de que sea cero en infinito e infinita en el centro, tal y como en el átomo de hidrógeno. Dada esta premisa, es válido proponer como una solución a la función

$$\mathbb{R}(\rho) = e^{-\frac{1}{2}\rho}G(\rho)$$

Si G fuera infinita sería una constante, como no lo es entonces es una función. Sustituyendo esta solución en la ecuación radial del átomo de hidrógeno se logra que

$$\frac{d^2G(\rho)}{d\rho^2} - \left(1 - \frac{2}{\rho}\right)\frac{dG(\rho)}{d\rho} + \left(\frac{\beta - 1}{\rho} - \frac{\ell(\ell+1)}{\rho^2}\right)G(\rho) = 0$$

La cual es de la forma $y'' + a(x)y' + b(x)y = 0$, y también cumple con tener singularidades en $\rho = 0$. En particular tiene una singularidad en el origen, ya que los coeficientes de $G\ y\ G'$ tienden a infinito cuando $\rho \longrightarrow 0$.

Ahora bien, como físicamente esperamos que la probabilidad de tener al electrón en el origen sea cero, entonces

$$G(\rho) = \rho L(\rho)$$

Donde L sea finita $L(\rho = 0) \neq 0$. Es decir, se puede proponer una solución del tipo

$$G(\rho) = \rho^s L(\rho)$$

Al sustituir esta solución se llega a

$$\rho^2 \frac{d^2L(\rho)}{d\rho^2} + \rho(2(s+1) - \rho)\frac{dL(\rho)}{d\rho} + [\rho(\beta - s - 1) + s(s+1) - \ell(\ell+1)]L(\rho) = 0$$

Para $\rho = 0$ se cuenta con

$$[s(s+1) - \ell(\ell+1)]\, L(0) = 0$$

$$\therefore \quad s(s+1) - \ell(\ell+1) = 0$$

$$s^2 + s - \ell(\ell+1) = 0$$

Tiene soluciones

$$\begin{cases} s_1 = \ell \\ s_2 = -\ell - 1 \end{cases}$$

Por lo tanto, hay dos soluciones

$$G(\rho) = \rho^{\ell} L(\rho)$$

$$G(\rho) = \frac{1}{\rho^{\ell+1}} L(\rho)$$

Esta última no es una solución físicamente aceptable, ya que para $\rho = 0$ se tiene que $G(\rho) \longrightarrow \infty$. Por lo tanto

$$\rho \frac{d^2L(\rho)}{d\rho^2} + (2(\ell+1) - \rho)\frac{dL(\rho)}{d\rho} + (\beta - \ell - 1)L(\rho) = 0$$

Las soluciones físicamente aceptables de esta ecuación con $\beta = n$ se pueden expresar en términos de los polinomios de Laguerre $L_q(\rho)$, que definen en términos de una función generadora:

$$U(\rho, s) = \frac{e^{-\frac{\rho s}{1-s}}}{1-s}$$

$$= \sum_{q=0}^{\infty} \frac{L_q(\rho)}{q!} s^q$$

Haciendo los siguientes cambios de parámetros en la ecuación diferencial anterior

$$p = 2\ell + 1$$

$$q = \beta + \ell$$

Se encuentra la ecuación asociada de Laguerre

$$\rho \frac{d^2L(\rho)}{d\rho^2} + (p + 1 - \rho)\frac{dL(\rho)}{d\rho} + (q - p)L(\rho) = 0$$

Así que a continuación seguiremos usando la ecuación anterior a esta. Al hacer un análisis de los puntos singulares $\rho = 0$, con singularidad regular:

$$L(\rho) = \sum_{n=0}^{\infty} \rho^n$$

$$\frac{dL(\rho)}{d\rho} = \sum_{n=0}^{\infty} n\rho^{n-1}$$

$$\frac{d^2L(\rho)}{d\rho^2} = \sum_{n=0}^{\infty} n(n-1)\rho^{n-2}$$

y sustituir se llega a

$$\rho \sum_{n=0}^{\infty} n(n-1)a_n\rho^{n-2} + (2(\ell+1)-\rho)\sum_{n=0}^{\infty} na_n\rho^{n-1} + (\beta-\ell-1)\sum_{n=0}^{\infty} a_n\rho^n = 0$$

Dividiendo entre ρ se tiene que

$$\sum_{n=0}^{\infty} n(n-1)a_n\rho^{n-2} + (2(\ell+1)-\rho)\sum_{n=0}^{\infty} na_n\rho^{n-2} + (\beta-\ell-1)\sum_{n=0}^{\infty} a_n\rho^{n-1} = 0$$

$$\sum_{n=0}^{\infty} n(n-1)a_n\rho^{n-2} + 2(\ell+1)\sum_{n=0}^{\infty} na_n\rho^{n-2} - \sum_{n=0}^{\infty} na_n\rho^{n-1} + (\beta-\ell-1)\sum_{n=0}^{\infty} a_n\rho^{n-1} = 0$$

Al factorizar se tiene

$$\sum_{n=0}^{\infty} \{n(n-1) + 2(\ell+1)n\}\, a_n\rho^{n-2} + \sum_{n=0}^{\infty} \{\beta-\ell-1-n\}a_n\rho^{n-1} = 0$$

tomando $n = N+1$ en la primera sumatoria, se encuentra

$$\sum_{N=-1}^{\infty} \{N(N+1) + 2(\ell+1)(N+1)\}\, a_{N+1}\rho^{N-1} + \sum_{n=0}^{\infty} \{\beta-\ell-1-n\}a_n\rho^{n-1} = 0$$

como para $N = -1$ se anula la sumatoria, entonces tomamos $N = 0$ y cambiamos de índice a $n = N$,

$$\therefore \quad \sum_{N=-1}^{\infty} \{n(n+1) + 2(\ell+1)(n+1)\}\, a_{n+1}\rho^{n-1} + \sum_{n=0}^{\infty} \{\beta-\ell-1-n\}a_n\rho^{n-1} = 0$$

de donde

$$\sum_{n=0}^{\infty} \{[n(n+1) + 2(\ell+1)(n+1)]a_{n+1} + [\beta-\ell-1-n]a_n\}\, \rho^{n-1} = 0$$

Esta igualdad se cumple si

$$[n(n+1)+2(\ell+1)(n+1)]a_{n+1}+[\beta-\ell-1-n]a_n=0$$

Por lo tanto, se tiene la relación de recurrencia

$$\frac{a_{n+1}}{a_n}=\frac{(\beta-\ell-1-n)}{(n+1)(n+2\ell+2)}$$

Un análisis del límite nos conduce a su convergencia (en este caso cero):

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\frac{n}{n^2}=\frac{1}{n}=0$$

Recordando que $\mathbb{R}(\rho)=e^{-\frac{1}{2}\rho}G(\rho)$ para el caso en que $G(\rho)=\rho^{\ell}L(\rho)$ se tiene que

$$\mathbb{R}(\rho)=\frac{\rho^{\ell}\sum_{n=0}^{\infty}a_n\rho^n}{e^{\frac{1}{2}\rho}}$$

Para observar que término domina, la serie de una función exponencial es $e^x=\sum_{n=0}^{\infty}\frac{1}{n!}x^n$, esto es:

$$e^{\frac{1}{2}\rho}=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{\rho}{2}\right)^n$$

$$=\sum_{n=0}^{\infty}\frac{1}{n!}\frac{1}{2^n}\rho^n$$

es decir,

$$e^{\frac{1}{2}\rho}=\sum_{n=0}^{\infty}a_n\,\rho^n \qquad donde \qquad a_n=\frac{1}{n!}\frac{1}{2^n}$$

En consecuencia

$$a_{n+1}=\frac{1}{(n+1)!}\frac{1}{2^{n+1}}$$

$$\Longrightarrow\frac{a_n}{a_{n+1}}=\frac{\frac{1}{n!}}{\frac{1}{(n+1)!}}\frac{\frac{1}{2^n}}{\frac{1}{2^{n+1}}}\Longrightarrow\frac{a_n}{a_{n+1}}=\frac{(n+1)!}{n!}\frac{2^{n+1}}{2^n}\qquad\Longrightarrow\qquad\frac{a_n}{a_{n+1}}=(n+1)2$$

$$\therefore\qquad\frac{a_{n+1}}{a_n}=\frac{1}{2(n+1)}$$

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\frac{1}{2n}$$

Entonces, la serie $\mathbb{R}(\rho) = \frac{\rho^{\ell}\sum_{n=0}^{\infty} a_n \rho^n}{e^{\frac{1}{2}\rho}}$ diverge, y se tiene que cortar ya que

$$\frac{\rho^{\ell}\frac{1}{n}}{\frac{1}{2n}}$$

Diverge por el dominio que tiene el término de arriba; es decir, el numerador de la solución radial $\mathbb{R}(\rho)$ puede dominar a la exponencial, a menos que $L(\rho)$ tenga un número finito de términos. Ahora bien, la serie se corta para un cierto n_r el cual satisface que

$$\beta = n_r + \ell + 1$$

Por lo tanto, β también es un número positivo. Definiendo $\beta \equiv n$ se encuentra que

$$n = n_r + \ell + 1$$

de donde $n \geq \ell + 1$ es el número cuántico principal, y como $\beta = \frac{Ze^2}{\hbar}\left(\frac{m_1}{2|E|}\right)^{1/2}$ entonces

$$n = \frac{Ze^2}{\hbar}\left(\frac{m_1}{2|E|}\right)^{1/2} \qquad \Longrightarrow \qquad n^2 = \frac{Z^2e^4}{\hbar^2}\frac{m_1}{2|E|}$$

Como $|E| = -E$ se tiene que

$$E = -\frac{1}{2}\frac{Z^2e^4m_1}{\hbar^2}\frac{1}{n^2}$$

Al definir

$$\alpha = \frac{e^2}{c\hbar} \qquad \rightarrow \qquad e^4 = \alpha^2c^2\hbar^2$$

Se encuentra que

$$E = -\frac{1}{2}m_1c^2\frac{(Z\alpha)^2}{n^2} \qquad \Longrightarrow \qquad E \sim \frac{Z^2}{n^2}$$

La ecuación asociada de Laguerre está dada por

$$x\frac{d^2L_N^K(x)}{dx^2} + (\mathrm{K} + 1 - x)\frac{dL_N^K(x)}{d\rho} + NL_N^K(x) = 0$$

cuya solución es

$$L_N^K(x) = \frac{e^x x^{-K}}{N!}\frac{d^N}{dx^N}(e^{-x}x^{N+K})$$

Nótese que $\rho = x,\ 2(\ell + 1) = \mathrm{K} + 1 \implies 2\ell + 1 = \mathrm{K}$, y $N = n - \ell - 1$ $dado\ que\ \ \beta = n$, por lo tanto, nuestra solución es:

$$L_{n-\ell-1}^{2\ell+1}(\rho) = \frac{e^{\rho}\rho^{-2\ell-1}}{(n-\ell-1)!}\frac{d^{n-\ell-1}}{d\rho^{n-\ell-1}}\left(e^{-\rho}\rho^{n+\ell}\right)$$

Si el polinomio es de grado $n-\ell-1$, es decir que la función radial tiene n_r nodos (ceros) y que habrá n_r-1 protuberancias. Si $\ell=n-1$ entonces $n_r=n-(n-1)-1=0$.
La diferenciación de la solución a los polinomios $L_q(\rho)$ de Laguerre p veces con respecto a ρ da la función generadora para los polinomios asociados de Laguerre

$$U(\rho,s)=\frac{(-s)^p e^{-\frac{\rho s}{1-s}}}{(1-s)^{p+1}}=\sum_{q=p}^{\infty}\frac{L_q^p(\rho)}{q!}s^q$$

Sustituyendo, para $p=2\ell+1$ y $q=\beta+\ell$, e intercambiando el orden de las sumatorias, se puede verificar que la siguiente expresión explícita satisface la función generadora de los polinomios asociados de Laguerre:

$$L_{n+\ell}^{2\ell+1}(\rho)=\sum_{k=0}^{n-\ell-1}(-1)^{k+1}\frac{[(n+\ell)!]^2\rho^k}{(n-\ell-1-k)!\,(2\ell+1+k)!\,k!}$$

Recapitulando, la solución de la ecuación radial viene dada por

$$R(\rho)=e^{-\frac{1}{2}\rho}G(\rho)$$

A su vez

$$G(\rho)=\rho^{\ell}L(\rho)$$

Entonces obtiene:

$$R(\rho)=e^{-\frac{1}{2}\rho}\rho^{\ell}L(\rho)$$

Por lo tanto:

$$R_{n\ell}(\rho)=e^{-\frac{1}{2}\rho}\rho^{\ell}L_{n-\ell-1}^{2\ell+1}(\rho)$$

y recordando que $\rho=\alpha r$ con $\alpha=\left(\frac{8m_1|E|}{\hbar^2}\right)^{1/2}$ implica $\rho=\left(\frac{8m_1|E|}{\hbar^2}\right)^{1/2}r$, y como $|E|=\frac{Z^2e^4m}{2\hbar^2n^2}$, entonces

$$\rho=\left(\frac{8m_1^2Z^2e^4}{2\hbar^4n^2}\right)^{1/2}r\Longrightarrow\rho=\frac{2m_1Ze^2}{\hbar^2n}r$$

$\alpha=\frac{2m_1Ze^2}{\hbar^2n}$. Definiendo el radio de Bohr $a_0=\frac{\hbar^2}{m_1e^2}$ se encuentra

$$\alpha=\frac{2Z}{na_0}$$

Ahora bien, para cumplir con la condición de ortonormalidad de la función radial

$$\psi(r,\theta,\varphi)=R_{n\ell}(r)T(\theta)F(\varphi)$$

como la ecuación es para un electrón, se debe cumplir que la densidad de probabilidad de encontrar al electrón en un elemento de volumen es igual a uno, esto es

$$\int \psi^*(r,\theta,\varphi)\psi(r,\theta,\varphi)dV = 1$$

$$\int \psi^*(r,\theta,\varphi)\psi(r,\theta,\varphi)r^2 sen\theta d\theta d\varphi dr = 1$$

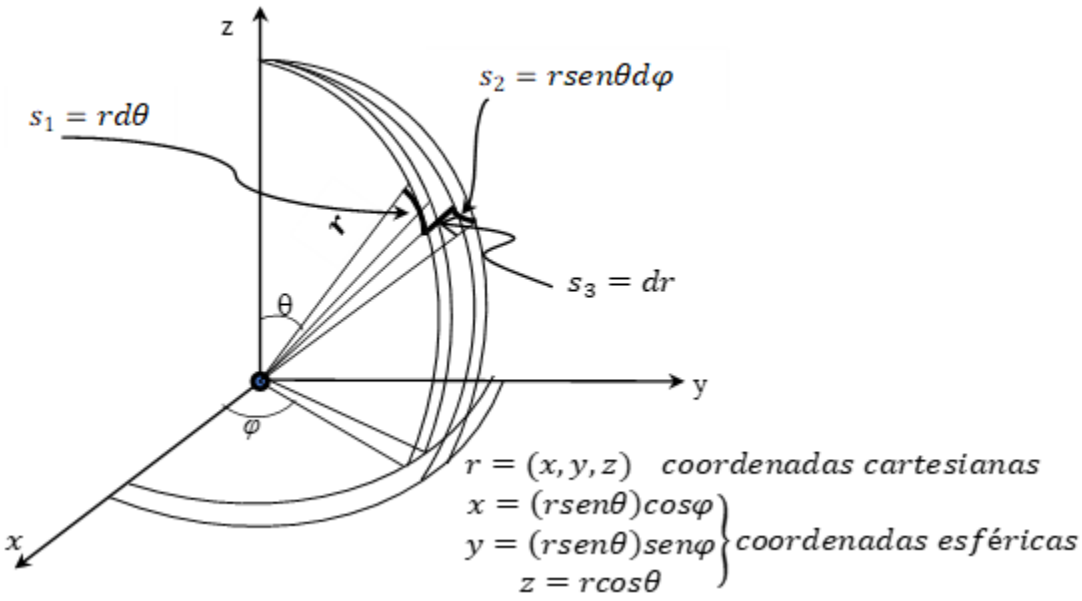
El elemento de volumen viene de las coordenadas esféricas, esto se consigue tomando un octete de esfera, como en la figura 11:

$$dV = s_1 s_2 s_3$$

$$dV = rd\theta(rsen\theta d\varphi)dr$$

$$\int_0^{\infty}\int_0^{2\pi}\int_0^{\pi} \psi^*(r,\theta,\varphi)\psi(r,\theta,\varphi)r^2 sen\theta d\theta d\varphi dr = 1$$

Figura 11. *Elemento de volumen*



Fuente: elaboración propia.

$$\int_0^{\infty} A^2 R_{n\ell}^*(r)R_{n\ell}(r)r^2 dr \int_0^{2\pi} B^2 T^*(\theta)T(\theta)sen\theta d\theta \int_0^{\pi} C^2 F^*(\varphi)F(\varphi)\, d\varphi = 1$$

Cada una de estas tres integrales es igual a 1, en particular

$$\int_0^{\infty} A^2\, R_{n\ell}^*(r)R_{n\ell}(r)r^2 dr = 1$$

La función de onda radial es de la forma $R_{n\ell}(\rho) = Ae^{-\frac{1}{2}\rho}\rho^{\ell}L_{n-\ell-1}^{2\ell+1}(\rho)$. La constante de normalización se puede encontrar usando la función generadora para evaluar la integral

$$\int_0^{\infty} e^{-\rho}(\rho)^{2\ell}\left[L_{n-\ell-1}^{2\ell+1}(\rho)\right]^2 \rho^2 d\rho = \frac{2n[(n+\ell)!]^3}{(n-\ell-1)!}$$

Entonces las funciones de onda radiales normalizadas del átomo de hidrógeno son

$$R_{n\ell}(\rho) = \left\{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-\ell-1)!}{2n[(n+\ell)!]^3}\right\}^{\frac{1}{2}} e^{-\frac{1}{2}\rho}\rho^{\ell}L_{n-\ell-1}^{2\ell+1}(\rho)$$

O bien $$R_{n\ell}(\rho) = -\left\{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-\ell-1)!}{2n[(n+\ell)!]^3}\right\}^{\frac{1}{2}} e^{-\frac{1}{2}\rho}\rho^{\ell}L_{n+\ell}^{2\ell+1}(\rho)$$

con $a_0 = \frac{\hbar^2}{me^2}$ (radio de Bohr) y $\rho = \frac{2Z}{na_0}r$. Las primeras tres funciones radiales son

$$R_{10}(r) = \left(\frac{Z}{a_0}\right)^{\frac{3}{2}} 2e^{-\frac{Zr}{a_0}} \qquad ya\ que\ \ L_0^1(\rho) = \frac{e^{\rho}\rho^{-1}}{1}\frac{d^0}{d\rho^0}(e^{-\rho}\rho^1) = 1$$

$$R_{20}(r) = \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}}\left(2 - \frac{Zr}{a_0}\right)e^{-\frac{Zr}{2a_0}} \quad ya\ que\ \ L_1^1(\rho) = \frac{e^{\rho}\rho^{-1}}{1}\frac{d}{d\rho}(e^{-\rho}\rho^2) = 2 - \rho$$

$$R_{21}(r) = \left\{\left(\frac{Z}{a_0}\right)^3 \frac{0!}{4(3!)^3}\right\}^{\frac{1}{2}} \frac{Zr}{a_0}e^{-\frac{Zr}{2a_0}}\left[\frac{(3!)^2}{3!}\right] = \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \frac{Zr}{a_0\sqrt{3}}e^{-\frac{Zr}{2a_0}}$$

La probabilidad de encontrar al electrón en una región limitada por las esferas de radio $r\ y\ r + dr$ es:

$$\psi^*(r,\theta,\varphi)\psi(r,\theta,\varphi)dV$$

para la parte radial se tiene

$$\mathrm{P}_{n\ell}(r) = R_{n\ell}^*(r)R_{n\ell}(r)r^2$$

a esta función se le conoce como densidad de probabilidad radial. La posición en la cual esta densidad es un máximo para el estado base del átomo de hidrógeno cuando $n = 1, \ell = 0$ es:

$$\mathrm{P}_{10}(r) = 2\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}} 2\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}} r^2 = 4\left(\frac{Z}{a_0}\right)^3 e^{-\frac{2Z}{a_0}r}$$

Para saber si tiene un mínimo o un máximo, derivamos

$$\frac{\mathrm{dP}_{10}(r)}{dr} = 4\left(\frac{Z}{a_0}\right)^3\left[\left(-\frac{2Z}{a_0}\right)e^{-\frac{2Z}{a_0}r}r^2 + e^{-\frac{2Z}{a_0}r}2r\right] = 0$$

$$= 4\left(\frac{Z}{a_0}\right)^3 e^{-\frac{2Z}{a_0}r}\left[-\frac{2Z}{a_0}r^2 + 2r\right] = 0$$

Como $= 4\left(\frac{Z}{a_0}\right)^3 e^{-\frac{2Z}{a_0}r}$ nunca es cero excepto en el infinito, entonces $-\frac{2Z}{a_0}r^2 + 2r = 0$. En este caso $Z = 1$. Por lo tanto $r = a_0$ es la posición en la cual la densidad de probabilidad es máxima. Ahora vamos a calcular el valor de expectación de la coordenada radial de un electrón en el átomo de hidrógeno cuya función radial viene dada por

$$\langle r_{n\ell}\rangle = \int_0^\infty r\, \mathrm{P}_{n\ell}(r) r^2 dr$$

$$= \int_0^\infty r^3\, R^*_{n\ell}(r) R_{n\ell}(r) dr = \frac{n^2 a_0}{Z}\left\{1 + \frac{1}{2}\left[1 + \frac{\ell(\ell+1)}{n^2}\right]\right\}$$

Para $n = 1$ se tiene que

$$\langle r_{n\ell}\rangle = 1.5 a_0$$

Recolectando todo, la solución al problema del átomo de hidrógeno es

$$\psi(r,\theta,\varphi) = R_{n\ell}(r) Y_\ell^m(\theta,\varphi)$$

$$\psi_{nml}(r,\theta,\varphi)$$

$$= \left\{\alpha^3 \frac{(n-\ell-1)!}{2n[(n+\ell)!]^3}\right\}^{\frac{1}{2}} e^{-\frac{1}{2}\alpha r}(\alpha r)^\ell L^{2\ell+1}_{n-\ell-1}(\alpha r)\left[\frac{2\ell+1}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}\right]^{1/2} P_\ell^{|m|}(cos\theta) e^{im\varphi}$$

Como $n = n_r + \ell + 1$ se tiene que $n \geq \ell + 1$. Además, como $\ell \geq |m|$ y ℓ es entero positivo, entonces $n \geq 1$ siempre. Por tanto, los números cuánticos están definidos como n principal, ℓ angular y m azimutal.

Si $n = 0$ entonces $\ell = 0 \implies m = 0$.

Si $n = 1$ entonces $\ell = 0 \implies m = 0$; $\ell = 1 \implies \begin{cases} m = 1 \\ m = 0 \\ m = -1 \end{cases}$

Si $n = 2$ entonces $\ell = 0 \implies m = 0$; $\ell = 1 \implies \begin{cases} m = 1 \\ m = 0 \\ m = -1 \end{cases}$; $\ell = 2 \implies \begin{cases} m = 2 \\ m = 1 \\ m = 0 \\ m = -1 \\ m = -2 \end{cases}$

Y así sucesivamente. Además,

$$\ell = 0 \longrightarrow s \qquad sharp$$
$$\ell = 1 \longrightarrow p \qquad principal$$
$$\ell = 2 \longrightarrow d \qquad diffuse$$

Dado que la energía del átomo de hidrógeno está cuantizada y restringida a los valores

$$E_n = -\frac{1}{2}\frac{Z^2 e^4 m_1}{\hbar^2}\frac{1}{n^2} \qquad n = 1,2,3,\ldots,$$

Degeneración

Ahora bien, si cada nivel energético o eigenvalor (valor propio) le corresponde a una sola eigenfunción (función propia), entonces el conjunto de funciones propias es *no degenerado*. Si para el *n-ésimo* estado propio existen g_n funciones propias entonces el *n-ésimo* estado propio es g_n veces *degenerado*.

Como los eigenvalores de la energía del átomo de hidrógeno dependen solo de n, entonces son degenerados con respecto a ℓ y m. Para cada valor de n, ℓ puede variar de 0 a $n-1$, y para cada uno de esos ℓ valores, m puede variar de $-\ell \ a \ +\ell$. La degeneración total de los niveles de energía E_n es, entonces

$$\sum_{\ell=0}^{n-1}(2\ell+1) = 2\frac{n(n-1)}{2} + n = n^2$$

La degeneración respecto a m es característica de cualquier campo de fuerzas centrales, para el cual V depende solo de la distancia radial r de algún punto. Sin embargo, la degeneración ℓ es característica del campo de Coulomb. Si se impone un campo magnético externo, la simetría esférica se destruye, la $(2\ell+1)$-veces degeneración desaparece, y el n-ésimo nivel hidrogenoide se desdobla en $2n^2$ niveles distintos. La degeneración ocurre cuando la ecuación de onda se puede resolver en más de una manera (en diferentes sistemas de coordenadas o en un solo sistema de coordenadas orientado de diferentes maneras), ya que si no hubiera degeneración las funciones de onda obtenidas en los diferentes sistemas de coordenadas tendrían que ser idénticas excepto por una constante multiplicativa, y esto usualmente no es posible. Para un campo central ocurre una excepción cuando $\ell = 0$, ya que la función de onda es esféricamente simétrica y tiene la misma forma para todas las orientaciones del eje polar, tal que no hay degeneración. En el problema del átomo de hidrógeno ocurre otra excepción similar cuando $n = 1$, en cuyo caso resulta que las soluciones obtenidas por separación de variables, tanto en coordenadas esféricas como en parabólicas de la ecuación de onda, son idénticas.

Momento angular

El momento angular $\boldsymbol{L}$ de una partícula de momento lineal $\boldsymbol{p}$ y vector de posición $\boldsymbol{r}$ viene dado por la expresión clásica:

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & k \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix}$$

de donde observamos que en mecánica clásica el momento angular es un vector. En mecánica cuántica el operador momento angular asociado a una partícula se da por

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x & y & z \\ \frac{\hbar}{i}\frac{\partial}{\partial x} & \frac{\hbar}{i}\frac{\partial}{\partial y} & \frac{\hbar}{i}\frac{\partial}{\partial z} \end{vmatrix}$$

Al expandir el determinante, las componentes (L_x, L_y, L_z) del momento angular son:

$$L_x = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$

$$L_y = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$$

$$L_z = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

Por tanto, el momento angular se puede escribir como

$$\boldsymbol{L} = \boldsymbol{i}L_x + \boldsymbol{j}L_y + \boldsymbol{k}L_z$$

Donde $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ son vectores unitarios ortogonales a lo largo de los ejes $x, y, z,$ respectivamente. Y el producto punto del vector momento angular nos conduce a la magnitud de su momento

$$\boldsymbol{L} \cdot \boldsymbol{L} = \hat{L}^2 = L_x^2 + L_y^2 + L_z^2$$

Los operadores L_x, L_y, L_z, y $\hat{L}^2$ son frecuentemente expresados en coordenadas esféricas. Empecemos estudiando al átomo de hidrógeno nuevamente, ahora por medio de operadores de momento angular. Las coordenadas naturales son las coordenadas esféricas

$$x = (rsen\theta)cos\phi$$
$$y = (rsen\theta)sen\phi$$
$$z = rcos\theta$$

Por la regla de la cadena

$$\frac{\partial}{\partial x} = sen\theta cos\phi \frac{\partial}{\partial r} + \frac{1}{r} cos\theta cos\phi \frac{\partial}{\partial \theta} - \frac{sen\phi}{rsen\theta} \frac{\partial}{\partial \phi}$$

$$\frac{\partial}{\partial y} = sen\theta sen\phi \frac{\partial}{\partial r} + \frac{1}{r} cos\theta sen\phi \frac{\partial}{\partial \theta} + \frac{cos\phi}{rsen\theta} \frac{\partial}{\partial \phi}$$

$$\frac{\partial}{\partial z} = cos\theta \frac{\partial}{\partial r} - \frac{sen\theta}{r} \frac{\partial}{\partial \theta}$$

de donde

$$L_x = \frac{\hbar}{i}\left(-sen\phi \frac{\partial}{\partial \theta} - cot\theta cos\phi \frac{\partial}{\partial \phi}\right)$$

$$L_y = \frac{\hbar}{i}\left(cos\phi \frac{\partial}{\partial \theta} - cot\theta sen\phi \frac{\partial}{\partial \phi}\right)$$

$$L_z = \frac{\hbar}{i} \frac{\partial}{\partial \phi}$$

$$\hat{L}^2 = -\hbar^2 \left(\frac{1}{sen\theta} \frac{\partial}{\partial \theta} sen\theta \frac{\partial}{\partial \theta} + \frac{1}{sen^2\theta} \frac{\partial^2}{\partial \phi^2}\right)$$

Ahora bien, recordando que la ecuación espacial de Schrödinger de un electrón en el átomo de hidrógeno viene dada por

$$-\frac{\hbar^2}{2m} \nabla_r^2 \mathcal{U}(\boldsymbol{r}) - \frac{Ze^2}{r} \mathcal{U}(\boldsymbol{r}) = E_r \mathcal{U}(\boldsymbol{r})$$

esto es $$\mathcal{H}\mathcal{U}(\boldsymbol{r}) = E_r \mathcal{U}(\boldsymbol{r})$$

donde el operador hamiltoniano es $\mathcal{H} = -\frac{\hbar^2}{2m} \nabla_r^2 + V(\boldsymbol{r})$ y al aplicar el operador a la función de onda $\psi = \psi(r, \theta, \phi)$ de Schrödinger se tiene que

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 sen^2\theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{r^2 sen\theta} \frac{\partial}{\partial \theta}\left(sen\theta \frac{\partial}{\partial \theta}\right)\right)\psi + V(r)\psi = E_r \psi$$

Esto se puede reescribir usando el operador de momento angular al cuadrado, es decir:

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2} \frac{1}{r^2} \hat{L}^2\right)\psi(r, \theta, \phi) + V(r)\psi(r, \theta, \phi) = E_r \psi(r, \theta, \phi)$$

Al tomar $\psi(r,\theta,\phi) = R(r)Y(\theta,\phi)$ y después de sustituir se encuentra que

$$\frac{r^2}{R(r)}\left(\frac{d^2R(r)}{dr^2}+\frac{2}{r}\frac{dR(r)}{dr}\right)+\frac{2mr^2}{\hbar^2}\left[E+\frac{q^2}{r}\right]=\ell(\ell+1)$$

Así la ecuación de eigenvalores para los armónicos esféricos ya va tomando forma, es decir

$$\therefore \qquad \hat{L}^2Y(\theta,\phi) = \ell(\ell\ +1)\ \hbar^2Y(\theta,\phi)$$

según los desarrollos anteriores. Esta última es una ecuación de eigenvalores del tipo

$$\hat{A}\psi = a\psi$$

Si separamos variables con $Y(\theta,\phi) = T(\theta)F(\phi)$ tenemos que

$$\frac{d^2F(\phi)}{d\phi^2} = -m^2F(\phi)$$

$$\frac{sen^2\theta}{T(\theta)}\frac{d^2T(\theta)}{d\phi^2}+cot\theta\frac{dT(\theta)}{d\theta}+\ell(\ell+1)-\frac{m^2}{sen^2\theta}=0$$

La ecuación diferencial con respecto a ϕ la podemos expresar en términos del operador L_z, esto es

$$\frac{i^2}{\hbar^2}\left(\frac{\hbar}{i}\frac{\partial}{\partial\phi}\left(\frac{\hbar}{i}\frac{\partial}{\partial\phi}\right)\right)F(\phi) = -m^2F(\phi)$$

o bien,

$$-\frac{1}{\hbar^2}\left(\hat{L}_z\hat{L}_z\right)F(\phi) = -m^2F(\phi)$$

$$\hat{L}_z^2F(\phi) = m^2\hbar^2F(\phi)$$

$$\hat{L}_zF(\phi) = m\hbar F(\phi)$$

Nótese que $\hat{L}_z, \hat{L}^2\ y\ \mathcal{H}$ son operadores que conmutan. Ahora consideramos el efecto de este operador sobre la función de onda

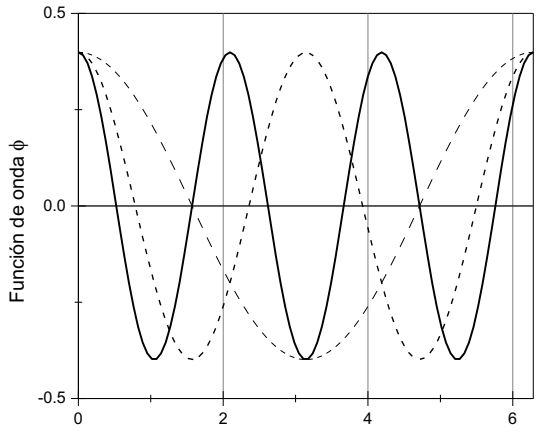
$$\hat{L}_z\phi_{m_\ell} = \frac{\hbar}{i}\frac{\partial}{\partial\phi}Ae^{im_\ell\phi} = m_\ell\hbar Ae^{im_\ell\phi} = m_\ell\hbar\phi_{m_\ell}$$

Aquí la función corresponde a un momento angular $m_l\hbar$. Si $m_\ell > 0$, por tanto el momento angular es positivo (al tomar el vector m_l con la mano derecha apuntando con el dedo pulgar en la dirección positiva de z, los demás dedos indican la dirección positiva de giro de este vector). Si $m_\ell < 0$, entonces el momento angular es negativo (ahora el dedo pulgar apunta en la dirección negativa del eje z y los demás dedos indican la dirección negativa de giro del vector). Las formas de las funciones de onda corresponden a un estado de momento angular definido como $m_\ell\hbar$:

$$\phi_{m_\ell} = \left(\frac{1}{2\pi}\right)^{1/2}e^{im_\ell\phi} = \left(\frac{1}{2\pi}\right)^{1/2}\{\cos m_\ell\phi + i\sin m_\ell\phi\}$$

Nótese que la función de onda es compleja. Si consideramos solo la parte real, cuando m_ℓ es entero, las funciones coseno forman una onda con un número entero de longitudes de onda. En la figura 12 se tienen una, dos y tres ondas cuando m_ℓ =1, 2 y 3, respectivamente, para una función de onda que satisface condiciones de frontera cíclicas.

Figura 12. *Función de onda con m_ℓ =1, 2 y 3 entre 0 y 2π*



Fuente: elaboración propia.

Los números cuánticos ℓ y m_l tienen un significado adicional. La energía de rotación de un cuerpo esférico de momento de inercia I y velocidad angular ω está dada por la física clásica como $E = \frac{1}{2} I\omega$. Debido a que la magnitud del momento angular se relaciona con la velocidad angular por $\ell = I\omega$, esta energía se puede expresar como $E = \ell^2/2I$. Usando la expresión anterior de la ecuación de eigenvalores para los armónicos esféricos se encuentra que:

$$magnitud\ del\ momento\ angular\ = \{\ell(\ell+1)\}^{1/2}\hbar$$

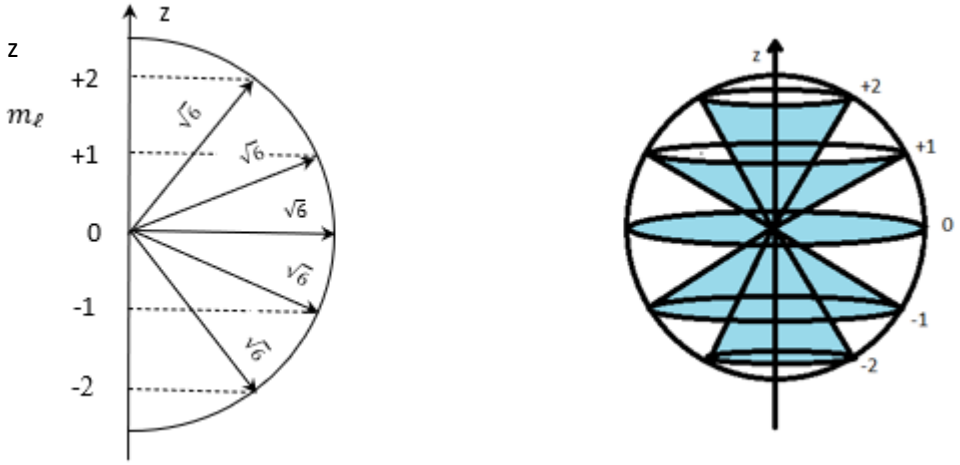
Así, la magnitud del momento angular queda cuantizada en mecánica cuántica. De hecho ℓ es llamado el número cuántico de momento angular. Los armónicos esféricos son también eigenfunciones de L_z:

$$L_z Y_{\ell m_\ell} = \frac{\hbar}{i}\frac{\partial}{\partial\phi}\left(\Theta_{\ell m_\ell}\frac{e^{\ell m_\ell \phi}}{\sqrt{2\pi}}\right) = m_\ell \hbar Y_{\ell m_\ell}$$

En este caso m_ℓ especifica la componente del momento angular alrededor del eje z. Sin embargo, debido a que m_ℓ está restringido a ciertos valores, la componente z del momento angular también está restringida a $2\ell + 1$ valores discretos para un valor de ℓ. Esta

restricción del componente de momento se llama *cuantización del espacio.* A m_ℓ se le conoce como momento magnético, ya que resulta de aplicar un campo magnético a un electrón en movimiento rectilíneo. En la figura 13 se muestran las cinco orientaciones permitidas del momento angular ($2\ell + 1$) con $\ell = 2$. La longitud del vector es $\{\ell(\ell+1)\}^{1/2} = 6^{1/2}$. Además, si la componente z del momento angular es especificada, las componentes x y y generalmente no pueden ser especificadas, el vector del momento angular podría estar mintiendo en una posición indeterminada.

Figura 13. *Componente z del momento angular*



Fuente: elaboración propia.

Los orbitales atómicos se identifican con los números cuánticos n, ℓ, m_ℓ, s. La configuración electrónica de los átomos depende del número de electrones que tengan. La tabla periódica se construye por el principio de Aufbau, respetando el principio de exclusión de Pauli y teniendo en cuenta la máxima multiplicidad o regla de Hund. Los orbitales en cada nivel energético van de la forma:

Nivel 1: contiene solo orbitales de tipo s

Nivel 2: contiene s *y* p

Nivel 3: s p d

Nivel 4: s p d f

Los orbitales s contienen a lo más 2 electrones; los p, 6 electrones; los d, 10 electrones; los f, 14 electrones. Primero hay que completar los orbitales s del nivel sucesivo

antes de completar los d, para que el átomo tenga una energía global lo más baja posible. Por ejemplo, el 4s se completa antes del 3d y el 5s antes del 4d (salvo excepciones). Las configuraciones electrónicas de los primeros átomos son las siguientes:

H: $1s^1$

He: $1s^2$

Li: $1s^2 2s^1$, o bien [He] $2s^1$

Be: $1s^2 2s^2$, o bien [He] $2s^2$

B: $1s^2 2s^2 p^1$, o bien [He] $2s^2 p^1$

C: $1s^2 2s^2 p^2$, o bien [He] $2s^2 p^2$

N: $1s^2 2s^2 p^3$, o bien [He] $2s^2 p^3$

O: $1s^2 2s^2 p^4$, o bien [He] $2s^2 p^4$

F: $1s^2 2s^2 p^5$, o bien [He] $2s^2 p^5$

Ne: $1s^2 2s^2 p^6$, o bien [He] $2s^2 p^6$

Na: $1s^2 2s^2 p^6 3s^1$, o bien [Ne] $3s^1$

Además de la configuración electrónica se debe considerar la multiplicidad. En general, por el acoplamiento espín-órbita que ocurre entre el espín resultante y el momento angular de un electrón, Russell-Saunders proponen el símbolo del término

$$^{2S+1}L_J$$

donde L es una letra (S, P, D, F, etc.) que corresponde al valor del número cuántico de momento angular orbital L (0, 1, 2, 3, etc.). La multiplicidad de un término es el valor $2S+1$, en la que S es el número cuántico del momento angular de espín total; como $L \geq S$ la multiplicidad es el número de niveles del término. Finalmente, J=L+S es el momento angular total. Realmente J = L+ S, L + S – 1, L + S – 2, … , |L – S|. El uso de las reglas de selección para construir estados de moléculas se puede consultar en Leighton (1959).

Método LCAO

Cada orbital molecular se expresa como una combinación lineal de orbitales atómicos (LCAO, por sus siglas en inglés) de varios átomos.

$$\psi_k = \sum_i c_{ik}\, \phi_i$$

donde ϕ_i es el orbital atómico i, y ψ_k es el orbital molecular k. Los orbitales forman un conjunto base y deben estar normalizados.

$$\psi_k = \int \phi_i \, \phi_i d\tau = 1$$

Con el método LCAO se forma la ecuación secular, que es una forma particular de la ecuación de onda de Schrödinger $\mathcal{H}\psi = E\psi$:

$$\mathcal{H}\psi - E\psi = (\mathcal{H} - E)\psi = 0$$

Se puede expresar como combinaciones lineales de orbitales atómicos

$$\sum_i c_i \, (\mathcal{H} - E)\phi_i = 0$$

En forma desarrollada para dos orbitales atómicos es $c_1(\mathcal{H} - E)\phi_1 + c_2(\mathcal{H} - E)\phi_2 = 0$, y multiplicando por ϕ_1 en un caso y por ϕ_2 en el otro caso se obtienen

$$c_1 \int \phi_1(\mathcal{H} - E)\phi_1 d\tau + c_2 \int \phi_1(\mathcal{H} - E)\phi_2 d\tau = 0$$

$$c_1 \int \phi_2(\mathcal{H} - E)\phi_1 d\tau + c_2 \int \phi_2(\mathcal{H} - E)\phi_2 d\tau = 0$$

de donde en general se definen

$$H_{ii} = \int \phi_i \mathcal{H} \phi_i d\tau$$

$$H_{ij} = \int \phi_i \mathcal{H} \phi_j d\tau$$

$$S_{ij} = \int \phi_i \phi_j d\tau$$

Las H_{ii} dan la energía del orbital atómico ϕ_i. Las H_{ij} dan las energías de interacción entre orbitales atómicos, y las S_{ij} dan el traslape o superposición. Aplicando estas integrales a nuestro sistema de dos orbitales se obtienen **H-S**E

$$c_1(H_{11} - E) + c_2(H_{12} - ES_{12}) = 0$$

$$c_1(H_{21} - ES_{21}) + c_2(H_{22} - E) = 0$$

o bien

$$\begin{pmatrix} H_{11} - E & H_{12} - ES_{12} \\ H_{21} - ES_{21} & H_{22} - E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0$$

$$(\mathbf{H} - \mathbf{S}E)\mathbf{c} = 0$$

Estas dos ecuaciones tienen soluciones no-triviales ($c_i = 0$ es trivial) cuando la matriz de los coeficientes forma un determinante igual a cero (*determinante secular*) $|\mathbf{H} - \mathbf{S}E| = 0$

$$\begin{vmatrix} H_{11} - E & H_{12} - ES_{12} \\ H_{21} - ES_{21} & H_{22} - E \end{vmatrix} = 0$$

de aquí se obtiene una ecuación polinomial cuadrática (*la ecuación secular*) cuyas energías corresponden al estado base y primer estado excitado.

Si los dos átomos son idénticos (molécula diatómica homonuclear) es aparente que $H_{11} = H_{22}$ y $H_{12} = H_{21}$. Aunque no pueda ser completamente obvio, $S_{12} = S_{21}$también, y estas integrales se llamarán simplemente *S*, por ahora. Entonces, resolviendo el determinante se tiene

$$(H_{11} - E)^2 - (H_{12} - SE)^2 = 0$$

Despejando se encuentra que

$$H_{11} - E = \pm(H_{12} - SE)$$

$$(1 \pm S)E = \pm H_{12} + H_{11}$$

O bien $$E_+ = \frac{H_{11} + H_{12}}{1 + S} \qquad E_- = \frac{H_{11} - H_{12}}{1 - S}$$

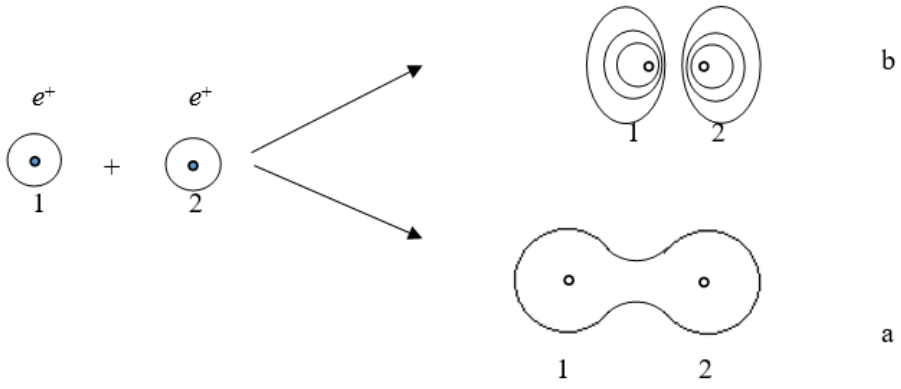
Para estos dos estados de energía, E_+ es el estado simétrico o enlazante, y E_- es el estado antisimétrico o antienlazante. Denotando los coeficientes asociados con E_+ como c_{11} y c_{12}, y los coeficientes asociados con E_- como c_{21} y c_{22}, entonces

$$\frac{c_{11}}{c_{12}} = +1 \qquad \frac{c_{21}}{c_{22}} = -1$$

$$\psi_1 = 1\sigma_g = N_1(1s_1 + 1s_2) \qquad \psi_2 = 1\sigma_u = N_2(1s_1 - 1s_2)$$

donde N_1 y N_2 son factores de normalización. Usando simetría, la función de onda ψ_1 del átomo 1 se denota como $1\sigma_g$ debido a que la combinación lineal de orbitales atómicos $(1s_1 + 1s_2)$ es *simétrica* (g viene de *gerade* en alemán, o 'par') bajo la operación de simetría de inversión, y la función de onda ψ_2 del átomo 2 se denota como $1\sigma_u$ debido a que la combinación lineal de orbitales atómicos $(1s_1 - 1s_2)$ es *antisimétrica* (u viene de *ungerade* en alemán, o impar) bajo inversión. La σ denota simetría cilíndrica de estas funciones alrededor del enlace. Las combinaciones de las funciones de onda se muestran gráficamente en la figura 14.

Figura 14. *Contornos de densidad electrónica de orbitales σ: a) enlazantes, b) antienlazantes*



Fuente: elaboración propia.

La combinación lineal de dos orbitales atómicos debido a la superposición de sus orbitales conduce a un orbital molecular diatómico, ya sea homonuclear o heteronuclear. Nótese que: i) dado que la ecuación de Schrödinger es lineal, si ψ_1 y ψ_2 son soluciones a esta entonces, su combinación lineal $c_1\psi_1 + c_2\psi_2$ también es solución, ii) dado que los operadores deben ser hermitianos, los eigenvalores son cantidades reales, iii) las eigenfunciones de operadores hermitianos son ortonormales: si $m = n$, $\int|\psi_m|^2 dx = 1$; y si $m \neq n$ y el sistema es no-degenerado $\int \psi_m{}^* \ \psi_n dx = 0$. La normalización es entonces:

$$\int|\psi_1|^2 d\tau = 1$$

$$\int|c_{11}(1s_1 + 1s_2)|^2 d\tau = 1$$

$$|c_{11}|^2 \int|1s_1 + 1s_2|^2 d\tau = 1$$

$$|c_{11}|^2 \int(|1s_1|^2 + |1s_2|^2 + |2s_1 s_2|)d\tau = 1$$

$$|c_{11}|^2(1 + 1 + 2S) = 1$$

$$c_{11} = \frac{1}{\{2(1+S)\}^{1/2}}$$

Análogamente:

$$\int |\psi_2|^2 d\tau = 1$$

$$c_{21} = \frac{1}{\{2(1-S)\}^{1/2}}$$

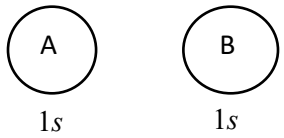
$$c_{11} = c_{12} \implies c_{11} = \frac{1}{\{2(1+S)\}^{1/2}} \implies para\ E_+ = \frac{H_{11} + H_{12}}{1+S}$$

$$c_{21} = c_{22} \implies c_{21} = \frac{1}{\{2(1-S)\}^{1/2}} \implies para\ E_- = \frac{H_{11} - H_{12}}{1-S}$$

Orbitales de moléculas diatómicas

Tomemos a continuación los átomos con menor número de electrones de la tabla periódica: dos átomos de hidrógeno aproximándose para formar una molécula (figura 15).

Figura 15. *Dos átomos de hidrógeno se aproximan para formar una molécula*



Fuente: elaboración propia.

Los electrones de los átomos de hidrógeno A y B pueden orbitar ambos núcleos si sus espines son opuestos. La función de onda constructiva molecular de 1s es la suma de dos orbitales atómicos 1s veces la constante (los orbitales de A y B tienen coordenadas diferentes):

$$\sigma_{1s} = c_1(\psi_{1s_A} + \psi_{1s_B})$$

La función de onda destructiva es la diferencia de dos orbitales atómicos 1s veces una constante

$$\sigma^*_{1s} = c_2(\psi_{1s_A} - \psi_{1s_B})$$

donde σ_{1s} y σ^*_{1s} deben ser ortogonales:

$$\psi_{1s_A} = \psi^*_{1s_A} = \frac{1}{\sqrt{\pi}}\left(\frac{1}{a_0}\right)^{3/2} e^{-\rho}$$

$$\psi_{1sB} = \psi^*_{1s_B} = \frac{1}{\sqrt{\pi}}\left(\frac{1}{a_0}\right)^{3/2} e^{-\rho}$$

ψ_{1s_A} y ψ_{1s_B} tienen las coordenadas diferentes.

$$\int \sigma_{1s}\sigma_{1s}^{*}dx = c_1c_2\int \left(\psi_{1s_A}+\psi_{1s_B}\right)\left(\psi_{1s_A}-\psi_{1s_B}\right)dx = c_1c_2\int \left[\left(\psi_{1s_A}\right)^2-\left(\psi_{1s_B}\right)^2\right]dx$$

$$= c_1c_2(1-1) = 0$$

Normalizando σ_{1s} se obtiene:

$$c_1^2\int \left(\psi_{1s_A}+\psi_{1s_B}\right)^2dx = c_1^2\left(2+2\int \psi_{1s_A}\cdot\psi_{1s_B}dx\right) = c_1^2(2+2\mathrm{S}) = 1$$

donde $\mathrm{S} = \int \psi_{1s_A}\cdot\psi_{1s_B}dx = \langle 1s_A|1s_B\rangle$, en notación de Dirac, es decir: $c_1 = \pm\frac{1}{\sqrt{2(1+\mathrm{S})}}$

$$\sigma_{1s} = \frac{1}{\sqrt{2(1+\mathrm{S})}}\left(\psi_{1s_A}+\psi_{1s_B}\right)$$

Si dos electrones son cercanos entre sí, $\psi_{1s_A}\approx\psi_{1s_B}$. La función de onda molecular constructiva $\sigma_{1s}\approx\frac{2\cdot\psi_{1s_A}}{\sqrt{2(1+\mathrm{S})}}$ indica mayor probabilidad de encontrar los electrones cercanos en la vecindad de cada núcleo. Se llama orbital enlazante porque una nube cargada negativamente entre los núcleos protegerá los núcleos positivos, y cada átomo puede formar una molécula. Normalizando σ_{1s}^{*} se tiene que:

$$\sigma_{1s}^{*} = \frac{1}{\sqrt{2(1-\mathrm{S})}}\left(\psi_{1s_A}-\psi_{1s_B}\right)$$

Si dos electrones son cercanos entre sí, la función de onda destructiva σ_{1s}^{*} tiende a cero. La probabilidad de encontrar los electrones es casi cero en σ_{1s}^{*}. Se llama orbital antienlazante ya que los núcleos positivos casi desnudos se enfrentan entre sí, y cada átomo no puede formar una molécula. Del mismo modo, se definieron los orbitales moleculares constructivo y destructivo de la combinación de dos orbitales atómicos 2s.

$$\sigma_{2s} = \frac{1}{\sqrt{2(1+\mathrm{S})}}\left(\psi_{1s_A}+\psi_{1s_B}\right)\qquad \sigma_{2s}^{*} = \frac{1}{\sqrt{2(1-\mathrm{S})}}\left(\psi_{1s_A}-\psi_{1s_B}\right)$$

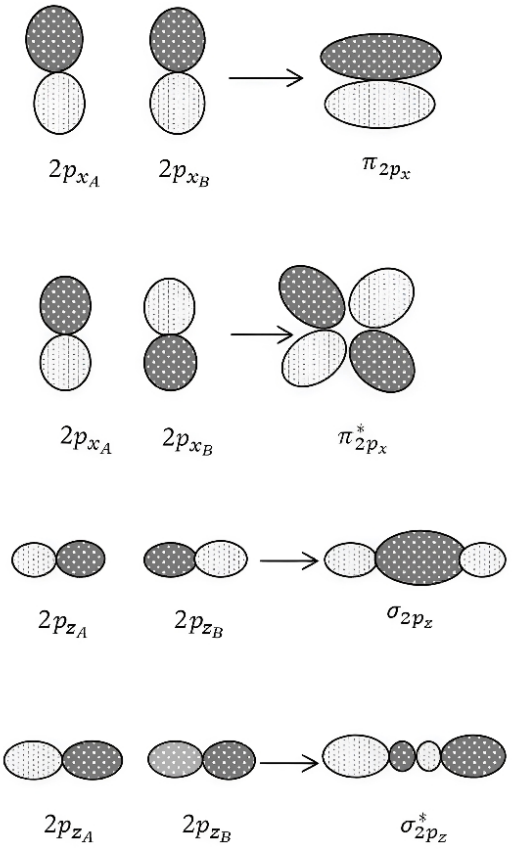
Orbitales enlazantes que se forman por combinación lineal de dos orbitales atómicos $2p_x$ o $2p_y$ (figura 16): $\pi-$ orbitales son orbitales moleculares constructivos de $2p_x$ o $2p_y$ con un plano nodal a través del eje de enlace.

π^*- orbitales resultantes de la interferencia destructiva de dos orbitales p_x o p_y y tienen un plano nodal perpendicular al eje de enlace internuclear. Tienen dos planos nodales.

Orbitales enlazantes que se forman por combinación lineal de dos orbitales atómicos de $2p_z$: Orbitales σ_{2p_z} resultantes de interferencia constructiva y no tienen planos nodales a lo largo del eje.

Orbitales $\sigma^*_{2p_z}$ provienen de interferencia destructiva y tienen un plano nodal perpendicular al eje de enlace perpendicular.

Figura 16. *Orbitales moleculares que originan p orbitales*



Fuente: elaboración propia.

Dos orbitales ortogonales tales como *s* y p_y, *s* y p_x, p_z y p_y, p_z y p_x, p_x y p_y no pueden traslapar para formar un orbital si fijamos el eje de enlace internuclear a z. Sin embargo, *s* y p_z pueden traslapar para formar un orbital.

Hidrógeno molecular diatómico H_2.

Configuración electrónica: $(\sigma_{1s})^2$

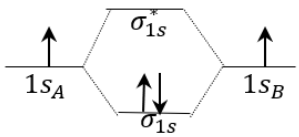
Electrones de valencia es: 1 + 1 = 2

Orden de enlace = $\frac{1}{2}$ electrones (enlace – antienlace)

= $\frac{1}{2}(2-0) = 1$. ∴ es enlace simple.

Energía de disociación: $E_d = 432\ KJ/mol$

H_2



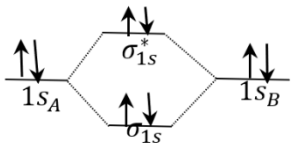
Helio molecular diatómico He_2

Configuración electrónica: $(\sigma_{1s})^2(\sigma_{1s}^*)^2$

Orden de enlace = 1/2 (2 – 2) = 0. Sin enlace. $E_d = 0.01\ KJ/mol$

He_2 tiene el enlace químico más débil conocido.

He_2



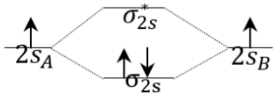
Litio Molecular diatómico Li_2

Configuración electrónica: $(\sigma_{1s})^2(\sigma_{1s}^*)^2(\sigma_{2s})^2$

Electrones de valencia es: 1 + 1 = 2

Orden de enlace =. 1/2 (2 – 0) = 1. $E_d = 105\ KJ/mol$

Li_2



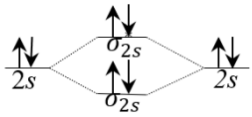
Berilio molecular diatómico Be_2

Configuración electrónica:$(\sigma_{1s})^2(\sigma_{1s}^*)^2(\sigma_{2s})^2(\sigma_{2s}^*)^2$

Electrones de valencia es: 2 + 2 = 4

Orden de enlace = 1/2 (2 – 2) = 0. $E_d = 9\ KJ/mol$

Be_2



Boro molecular diatómico B_2

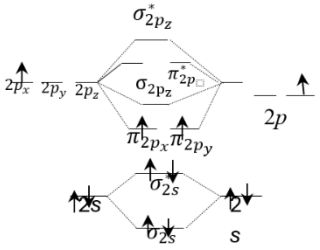
Configuración electrónica:

$$(\sigma_{1s})^2(\sigma_{1s}^*)^2(\sigma_{2s})^2(\sigma_{2s}^*)^2\left(\pi_{2p_x}\right)^1\left(\pi_{2p_y}\right)^1$$

Electrones de valencia: 3 + 3 = 6

Orden de enlace = 1/2 (4 – 2) = 1. $E_d = 289\ KJ/mol$

B_2



Carbono molecular diatómico C_2

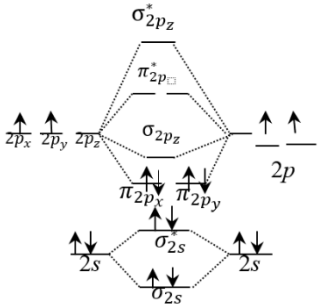
Configuración electrónica:

$$(\sigma_{1s})^2(\sigma_{1s}^*)^2(\sigma_{2s})^2(\sigma_{2s}^*)^2\left(\pi_{2p_x}\right)^2\left(\pi_{2p_y}\right)^2$$

Electrones de valencia: 4 + 4 = 8

Orden de enlace = 1/2 (6 – 2) = 2. $E_d = 599\ KJ/mol$

C_2



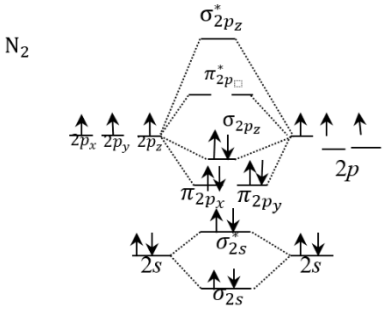
Nitrógeno molecular diatómico N_2

Configuración electrónica:

$$(\sigma_{1s})^2(\sigma_{1s}^*)^2(\sigma_{2s})^2(\sigma_{2s}^*)^2\left(\pi_{2p_x}\right)^2\left(\pi_{2p_y}\right)^2\left(\sigma_{2p_z}\right)^2$$

Electrones de valencia: 5 + 5 = 10

Orden de enlace = 1/2 (8 – 2) = 3. $\Delta E_d = 941\ KJ/mol$



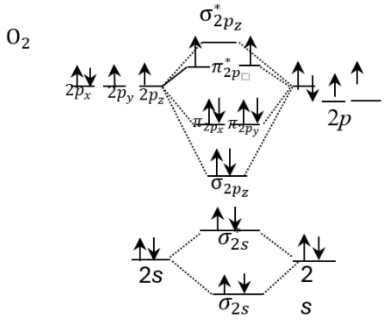
Oxígeno molecular diatómico O_2

Configuración electrónica:

$$(\sigma_{1s})^2(\sigma_{1s}^*)^2(\sigma_{2s})^2(\sigma_{2s}^*)^2\left(\sigma_{2p_z}\right)^2\left(\pi_{2p_x}\right)^2\left(\pi_{2p_y}\right)^2\left(\pi_{2p_x}^*\right)^1\left(\pi_{2p_y}^*\right)^1$$

Electrones de valencia: 6 + 6 = 12

Orden de enlace = 1/2 (8 – 4) = 2. $\Delta E_d = 494\ KJ/mol$



Hay dos electrones desapareados en orbitales π_{2p}^*. La energía relativa del orbital σ_{2p_z} comparada a la de los orbitales π_{2p} depende del valor Z de los átomos. Entonces, si Z ≥ 8, el orbital σ_{2p_z} es más bajo en energía. Como se puede observar, la energía de disociación incrementa conforme aumenta el orden de enlace. La energía de disociación es igual a la energía de enlace.

El estado o configuración electrónica de un átomo se denota con las letras de los orbitales del momento angular, es decir s, p, d, f. Cuando se tienen moléculas estos orbitales interaccionan para formar nuevos enlaces; las interacciones con enlaces s se convierten en un enlace σ, mientras que aquellas de enlace p se convierten en π, como se ha visto en las moléculas diatómicas anteriores. Aquellos de enlace d se convierten en δ, y así sucesivamente.

Orbitales de moléculas poliatómicas

Las moléculas diatómicas contemplan dos átomos ligados. Las moléculas con tres átomos ligados son moléculas triatómicas, aquellas con cuatro átomos ligados son tetraatómicas y así sucesivamente. En esta secuencia, todas las moléculas con más de dos átomos ligados corresponden a moléculas poliatómicas. Cuando existen interacciones entre enlaces que no son del mismo tipo se obtienen hibridaciones. A continuación se muestran hibridaciones de

orbitales atómicos de algunos casos simples e ilustrativos. El análisis de las moléculas poliatómicas se logra considerando su simetría, esto se ejemplifica con el grupo de simetría C_{3v}.

Hibridación de los orbitales atómicos

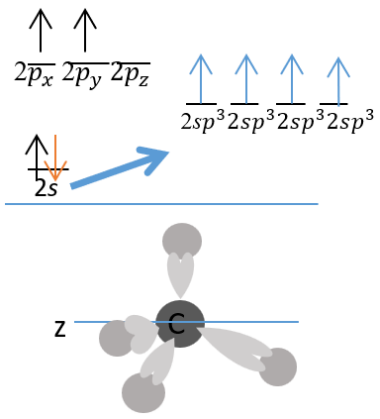
Para explicar el enlace en moléculas como el metano, debemos pensar en la hibridación de los orbitales s y p. Dado que el metano es una molécula en la que un carbono se une con cuatro hidrógenos en lugar de dos, la simple teoría del enlace de valencia no se puede aplicar para describir el enlace con precisión.

Suponemos que los orbitales atómicos cercanos en energía pueden hibridar (Atkins et al., 2008; Berkowski, 2005; Drennan y Taylor, 2008; Giancoli, 2009; McMurry, 2004). Los enlaces resultan del apareamiento o desapareamiento de electrones en orbitales atómicos o hibridados. Dos orbitales ortogonales tales como s y p_y, s y p_x, p_z y p_y, p_z y p_x, p_x y p_y no pueden superponerse para formar un orbital si fijamos el eje internuclear del átomo en z. Hay dos tipos de enlaces: σ y π. El enlace σ es cilíndricamente simétrico sin un plano nodal a través del eje del enlace: s y s, s y p_z, p_z y p_z , sp^3 y sp^3; s y sp^3 son los que forman los enlaces σ. El enlace π es un enlace con densidad de electrones en dos lóbulos a cada lado del eje de enlace: p_x y p_x, p_y y p_y son los que hacen el enlace π.

Hibridación sp^3

En un átomo de carbono tenemos el orbital-$2p_z$ vacío. Si un electrón de orbital-2s se promueve al orbital-$2p_z$ vacío, el átomo de carbono tendrá cuatro electrones no apareados listos para enlazarse. Los orbitales híbridos sp^3 son degenerados, solo difieren en su orientación en el espacio y se obtienen combinando un orbital-s y tres orbitales-p. Los lóbulos grandes de los orbitales hibridados están orientados de forma tetraédrica con un ángulo de enlace de 109.5° como resultado de la minimización de la repulsión

Figura 17. *Metano con cuatro enlaces σ*



Fuente: elaboración propia.

de electrones. El metano es un ejemplo de una molécula con orbitales híbridos-sp^3, en la cual el carbón tiene cuatro orbitales híbridos equivalentes participando en enlaces (figura 17). Estos cuatro orbitales son híbridos a partir de uno 2s y tres 2p, y se puede considerar cada uno de ellos como una combinación de $\frac{1}{2}s$ y $\frac{3}{4}p$, y a esta combinación es a la que se le llama híbrido-sp^3. La hibridación da como resultado una mayor liberación de energía, con lo cual la molécula resultante es más estable.

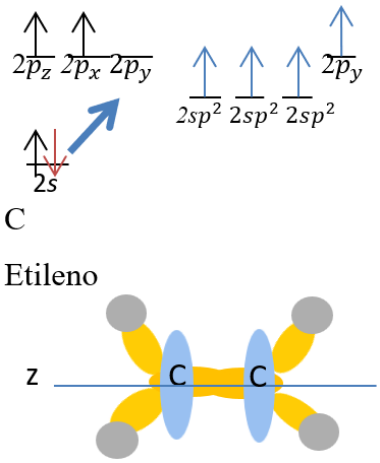
Si consideramos la geometría de NH_3 y el tipo de enlace N-H, no tenemos orbitales 2p vacíos. La promoción de electrones no puede ocurrir porque la promoción no aumenta el número de electrones no pareados para el enlace. Sin embargo, puede ocurrir que la hibridación tenga cuatro orbitales sp^3 al combinar orbitales 2s y tres orbitales 2p, lo que explicaría la geometría piramidal trigonal de NH_3. Si la hibridación no ocurriera, el ángulo de enlace sobre el átomo central N se convertiría en 90°, lo que contradice la teoría VSEPR (*valence shell electron pair repulsion*). El número total de electrones de valencia es 3 + 5 = 8. El número de electrones de enlace es 6. El número de electrones sin enlace es 8 - 6 = 2. El tipo de enlace N-H es σ (C $2sp^3$, H 1s). El ángulo de enlace es menor que 109.5 debido a un par solitario en N.

Hibridación sp^2

Orbitales sp^2 híbridos se forman al combinar un orbital-s y dos orbitales-p. El Boro es capaz de tener tres orbitales desapareados si un electrón-2s es promovido a un orbital-2p vacío e hibridiza. Los tres orbitales-sp^2 están en un plano para minimizar la repulsión electrónica. La geometría resultante es planar trigonal. El ángulo de enlace alrededor del átomo central es 120°. Un orbital-2p vacío aún está ahí, y disponible para aceptar electrones. La molécula BH_3 es un ejemplo de este tipo de enlace. El tipo de enlace B-H es $\sigma(\text{B } 2sp^2, \text{H } 1s)$.

El carbón también forma orbitales sp^2 híbridos. Como un resultado el carbón es capaz de tener un electrón desapareado en cada uno de los tres orbitales-$2sp^2$ y un orbital-2p. Dos orbitales-2 p_x o dos orbitales-2 p_y pueden traslapar constructivamente para formar un enlace π. Orbitales-2s p^2 pueden formar enlace σ con orbital-s u otro orbital-2s p^2. Dos orbitales

Figura 18. *Etileno con un enlace π en azul, y tres enlaces σ en amarillo*



Fuente: elaboración propia.

ortogonales, tales como s y p_y, s y p_x, no pueden traslapar para formar un orbital si fijamos el eje internuclear de C a z.

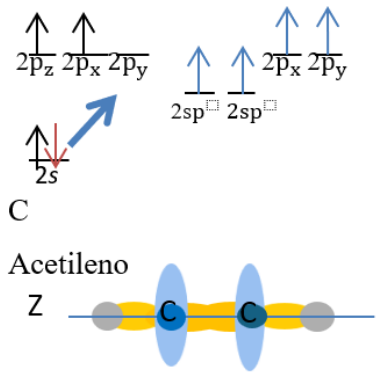
Etileno tiene un doble enlace C-C que consiste en un enlace $\sigma(\mathrm{C}\ 2sp^2, \mathrm{C}\ 2sp^2)$ y otro π(C $2p$, C $2p$). En adición con los dobles enlaces C-C, hay cuatro enlaces $\sigma(\mathrm{C}\ 2sp^2, \mathrm{H}\ 1s)$. Las moléculas no pueden rotar alrededor del doble enlace (figura 18).

Hibridación sp[1]

El acetileno es lineal y solo tiene un átomo de hidrógeno en cada extremo (como se ve en la figura 19).

Los orbitales híbridos sp^1 se forman al combinar un orbital que combina un orbital-s y un orbital-p. El carbón ahora puede tener un electrón desapareado en cada uno de los dos orbitales-sp. En total son cuatro electrones desapareados. El enlace C-C consiste en un enlace $\sigma(\mathrm{C}\ 2sp, \mathrm{C}\ 2sp)$ otro enlace $\pi(\mathrm{C}\ 2p_x, \mathrm{C}\ 2p_x)$ y otro $\pi\left(\mathrm{C}\ 2p_y, \mathrm{C}\ 2p_y\right)$ si el eje internuclear de C es z.

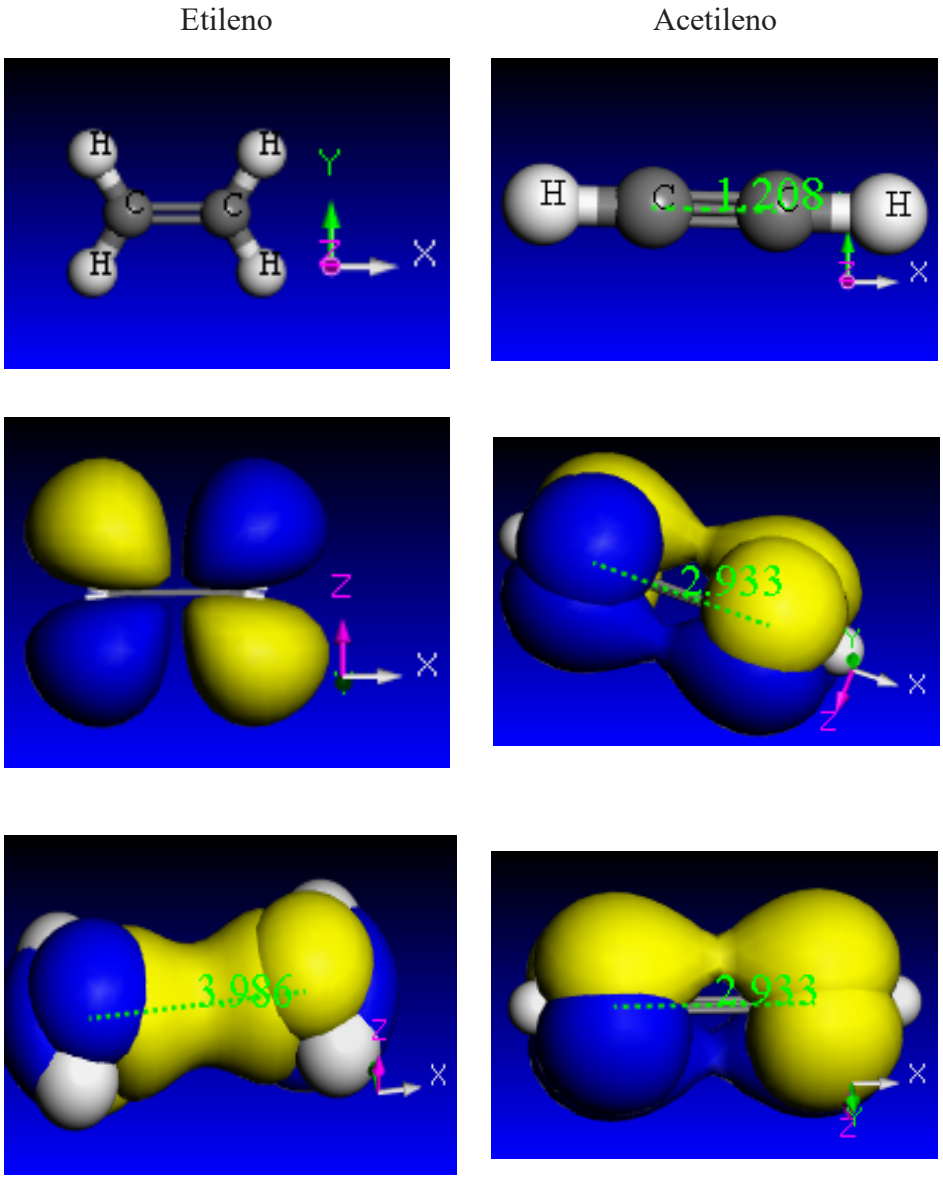
Figura 19. *Acetileno con dos enlaces π en azul y tres enlaces σ en amarillo*



Fuente: elaboración propia.

Las Figuras 18 y 19 para etileno y acetileno, adicionalmente se elaboraron usando BIOVIA Materials Studio 2017, como se ve en la figura 20.

Figura 20. *Orbitales de etileno (sp^2) y acetileno (sp^1) con sus respectivos enlaces σ y π*



Fuente: elaboración propia.

La multiplicidad en una molécula viene a partir de las bandas de estructura electrónica fina, análoga al caso de átomos, la estructura multiplete se debe al espín del electrón. El momento angular electrónico total alrededor del eje internuclear, denotado por Ω, se obtiene sumando Λ y Σ, tal y como el momento angular electrónico J se logra sumando L y S. Sin embargo, mientras que para átomos se debe llevar a cabo una adición vectorial, para moléculas es suficiente una adición algebraica, puesto que Λ y Σ están a lo largo de la línea

que une el núcleo. Entonces, para el número cuántico del momento angular electrónico resultante, alrededor del eje internuclear se tiene

$$\Omega = |\Lambda + \Sigma|$$

El vector Λ correspondiente de momento angular representa la componente del momento angular orbital electrónico a lo largo del eje internuclear, Λ = 0, 1, 2, … , L. De modo que en la molécula por cada valor de L existen L + 1 estados distintos con diferente energía. Los estados moleculares correspondientes se designan como Λ = Σ, Π, Δ, Φ, …

Los estados Π, Δ, Φ, … son doblemente degenerados puesto que M_L puede tomar los dos valores +Λ y -Λ; los estados Σ son no-degenerados.

$$\Sigma = S, S-1, S-2, \ldots, -S$$

Son 2S+1 valores diferentes posibles. En contraste a Λ, los números cuánticos Σ pueden ser positivos y negativos. No está definido para estados con Λ=0 – esto es estados Σ.

Simetrías

Los métodos mecánico-cuánticos para tratar la estructura electrónica de moléculas poliatómicas son los mismos que aquellos aplicados a las moléculas diatómicas. Un papel principal para encontrar la forma de las LCAO-MO en moléculas poliatómicas es el uso de la simetría. La teoría de grupos es la que nos permite trabajar con las simetrías de las moléculas. Las operaciones de simetría elementales son las siguientes:

E Identidad. El elemento de simetría es la operación del objeto mismo

C_n Rotación n-veces de. El elemento es la operación de rotación 2π/n alrededor de un eje de simetría.

σ Reflexión. El elemento es la operación en un espejo plano.

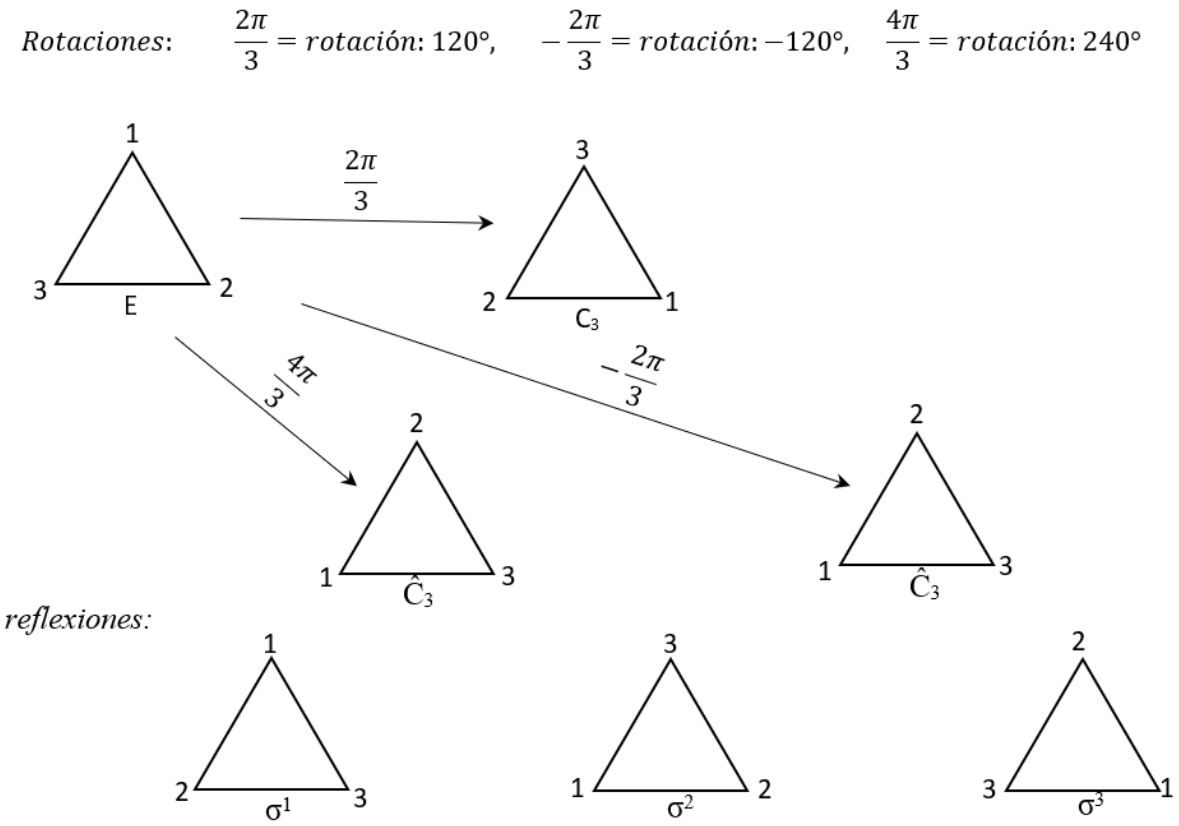
i Inversión. El elemento es la operación a través de un centro de simetría.

S_n Rotación impropia n-veces (reflexión-rotatoria). La operación ocurre alrededor de un eje de rotación impropia

En el sistema Schoenflies la notación permite sin confusión el uso de los grupos: C_1, C_s, C_i, C_{nv}, $C_{\infty v}$, C_{nh}, D_{nh}, $D_{\infty h}$, D_{nd}, I, I_h, O, O_h, R_3, T, T_d, T_h.

Como ejemplo se ilustra la tabla de multiplicación del grupo C_{3v}, el cual corresponde a seis elementos de simetría: identidad (E), dos rotaciones del triángulo (C_3, $\hat{C}_3$), tres

reflexiones (σ^1, σ^2, σ^3). A partir de operaciones de simetría con base en el elemento identidad (E) se construyen los demás elementos de la tabla.

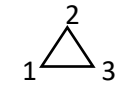


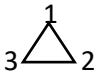
Estas son reflexiones σ^i a través del plano vertical al eje *i* y perpendicular al plano del triángulo. Las reflexiones se entienden como una permutación o intercambio de números respecto al elemento identidad E: $\sigma^1 = (23)$, $\sigma^2 = (31)$, $\sigma^3 = (21)$, de esta manera se van a hacer los productos para elaborar la tabla

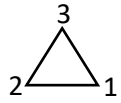
Operaciones

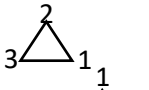
Las operaciones consisten en productos directos entre elementos del grupo, como ejemplos se tienen (R = rotación, r = reflexión: $RE = ER = R$; $RR^{-1} = R^{-1}R = E$; etc., comúnmente una operación R seguida de una operación r se denota por rR, y se enfatiza que el orden de la operación es importante, ya que en general la operación rR no es lo mismo que la operación Rr; esto es que no se cumple la conmutatividad, sin embargo, sí se cumple la asociatividad.

$EC_3 = C_3$

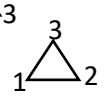
$C_3C_3 = \overset{\frown}{\frac{2\pi}{3}}\ \overset{\frown}{\frac{2\pi}{3}} = \hat{C}_3$ 

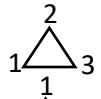
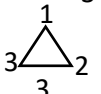
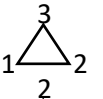
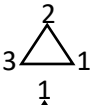
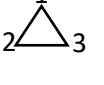
$\hat{C}_3C_3 = \overset{\curvearrowleft}{\frac{2\pi}{3}}\ \overset{\frown}{\frac{2\pi}{3}} = E$ 

C_3 

$\sigma^1C_3 = (23)\overset{\frown}{\frac{2\pi}{3}} = \sigma^3$ 

$\sigma^2C_3 = (31)\overset{\frown}{\frac{2\pi}{3}} = \sigma^1$ 

$\sigma^3C_3 = (21)\overset{\frown}{\frac{2\pi}{3}} = \sigma^2$ 

$C_3E = C_3$	$\hat{C}_3E = \hat{C}_3$	$\sigma^1E = \sigma^1$	$\sigma^2E = \sigma^2$	$\sigma^3E = \sigma^3$	
$C_3C_3 = \overset{\frown}{\frac{2\pi}{3}}\ \overset{\frown}{\frac{2\pi}{3}} = \hat{C}_3$ 	$\hat{C}_3C_3 = \overset{\curvearrowleft}{\frac{2\pi}{3}}\ \overset{\frown}{\frac{2\pi}{3}} = E$	$\sigma^1C_3 = \sigma^3$	$\sigma^2C_3 = \sigma^1$	$\sigma^3C_3 = \sigma^2$	
$C_3\hat{C}_3 = \overset{\frown}{\frac{2\pi}{3}}\ \overset{\curvearrowleft}{\frac{2\pi}{3}} = E$ 	$\hat{C}_3\hat{C}_3 = \overset{\curvearrowleft}{\frac{2\pi}{3}}\ \overset{\curvearrowleft}{\frac{2\pi}{3}} = C_3$	$\sigma^1\hat{C}_3 = \sigma^2$	$\sigma^2\hat{C}_3 = \sigma^3$	$\sigma^3\hat{C}_3 = \sigma^1$	
$C_3\sigma^1 = \overset{\frown}{\frac{2\pi}{3}}(23) = \sigma^2$ 	$\hat{C}_3\sigma^1 = \overset{\curvearrowleft}{\frac{2\pi}{3}}(23) = \sigma^3$	$\sigma^1\sigma^1 = E$	$\sigma^2\sigma^1 = C_3$	$\sigma^3\sigma^1 = \hat{C}_3$	
$C_3\sigma^2 = \overset{\frown}{\frac{2\pi}{3}}(31) = \sigma^3$ 	$\hat{C}_3\sigma^2 = \overset{\curvearrowleft}{\frac{2\pi}{3}}(31) = \sigma^1$	$\sigma^1\sigma^2 = \hat{C}_3$	$\sigma^2\sigma^2 = E$	$\sigma^3\sigma^2 = C_3$	
$C_3\sigma^3 = \overset{\frown}{\frac{2\pi}{3}}(21) = \sigma^1$ 	$\hat{C}_3\sigma^3 = \overset{\curvearrowleft}{\frac{2\pi}{3}}(21) = \sigma^2$	$\sigma^1\sigma^3 = C_3$	$\sigma^2\sigma^3 = \hat{C}_3$	$\sigma^3\sigma^3 = E$	

La flecha curva indica el giro del triángulo. La tabla de multiplicación queda del siguiente modo:

C_{3v}	E	C_3	$\hat{C}_3$	σ^1	σ^2	σ^3
E	E	C_3	$\hat{C}_3$	σ^1	σ^2	σ^3
C_3	C_3	$\hat{C}_3$	E	σ^2	σ^3	σ^1
$\hat{C}_3$	$\hat{C}_3$	E	C_3	σ^3	σ^1	σ^2
σ^1	σ^1	σ^3	σ^2	E	$\hat{C}_3$	C_3
σ^2	σ^2	σ^1	σ^3	C_3	E	$\hat{C}_3$
σ^3	σ^3	σ^2	σ^1	$\hat{C}_3$	C_3	E

En un tratamiento matricial alternativo se construyen por inspección los elementos de simetría del grupo C_{3v} usando una base propuesta con respecto a la molécula triangular (de lados de igual tamaño) piramidal NH_3, de la cual cada uno de sus átomos se considera de simetría esférica: s_n, s_a, s_b, s_c, por lo cual la base es (s_n, s_a, s_b, s_c), que bajo la operación σ_v se convierte en $\sigma_v(s_n, s_a, s_b, s_c) = (s_n, s_a, s_c, s_b)$. La transformación correspondiente se representa como la multiplicación matricial.

$$\sigma_{\mathrm{v}}(s_{\mathrm{n}}, s_{\mathrm{a}}, s_{\mathrm{b}}, s_{\mathrm{c}}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

La matriz representativa de esta operación se denota como $D(\sigma_{\mathrm{v}})$. Análogamente se construyen las demás, de tal modo que la representación matricial es:

$$D(E) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad D(C_3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad D\left(\hat{C}_3\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$D(\sigma_{\mathrm{v}}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad D(\sigma'_{\mathrm{v}}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad D(\sigma''_{\mathrm{v}}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

En general, las operaciones de simetría de la tabla de multiplicación del grupo C_{3v} se reproducen con productos entre estas matrices, dado que $RS=T$ es equivalente a $D(R)D(S)=D(T)$. Los elementos distintos de cero en la diagonal corresponden a $\chi(R) = trD(R)$; es decir: $\chi(E) = 4, \chi(C_3) = 1, \chi\left(\hat{C}_3\right) = 1,\ \chi(\sigma_{\mathrm{v}}) = 2, \chi(\sigma'_{\mathrm{v}}) = 2, \chi(\sigma''_{\mathrm{v}}) = 2.$

Con otra base se construye una transformación de semejanza. Para esto se elige la base $(s_{\mathrm{n}}, s_1, s_2, s_3)$ con la relación $s_{\mathrm{n}} = s_{\mathrm{n}}\,;\ s_1 = s_{\mathrm{a}} + s_{\mathrm{b}} + s_{\mathrm{c}}\,;\ s_2 = 2s_{\mathrm{a}} - s_{\mathrm{b}} - s_{\mathrm{c}}\,;\ s_3 = s_{\mathrm{b}} - s_{\mathrm{c}}$. En forma matricial estas relaciones se expresan como

$$(s_{\mathrm{n}}, s_1, s_2, s_3) = (s_{\mathrm{n}}, s_{\mathrm{a}}, s_{\mathrm{b}}, s_{\mathrm{c}}) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix} = (s_{\mathrm{n}}, s_{\mathrm{a}}, s_{\mathrm{b}}, s_{\mathrm{c}})c$$

La inversa de la matriz c es: $c^{-1} = \frac{1}{6}\begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 2 & 2 \\ 0 & 2 & -1 & -1 \\ 0 & 0 & 3 & -3 \end{bmatrix}$. La matriz representativa de C_3 en la nueva base es $D'(C_3) = c^{-1}D(C_3)c = \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 3 & -1 \end{bmatrix}$

Análogamente se construyen las demás transformaciones de semejanza, de manera que la representación matricial del grupo C_{3v} es la siguiente:

$$D(E)=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}, D(C_3)=\frac{1}{2}\begin{bmatrix}2&0&0&0\\0&2&0&0\\0&0&-1&-1\\0&0&3&-1\end{bmatrix},\ D\left(\hat{C}_3\right)=\frac{1}{2}\begin{bmatrix}2&0&0&0\\0&2&0&0\\0&0&-1&1\\0&0&-3&-1\end{bmatrix},$$

$$D(\sigma_{\mathrm{v}})=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix},\qquad D(\sigma_{\mathrm{v}}')=\frac{1}{2}\begin{bmatrix}2&0&0&0\\0&2&0&0\\0&0&-1&1\\0&0&3&1\end{bmatrix},\qquad D(\sigma_{\mathrm{v}}'')=\frac{1}{2}\begin{bmatrix}2&0&0&0\\0&2&0&0\\0&0&-1&-1\\0&0&-3&1\end{bmatrix}$$

La suma de los elementos en las diagonales es invariante en esta nueva representación del grupo C_{3v}, es decir que $\chi(E)=4, \chi(C_3)=1, \chi\left(\hat{C}_3\right)=1,\quad \chi(\sigma_{\mathrm{v}})=2,\quad \chi(\sigma_{\mathrm{v}}')=2, \chi(\sigma_{\mathrm{v}}'')=2$.

Todas las matrices de la representación matricial del grupo C_{3v} tienen una forma de bloque-diagonal. La base cuatro-dimensional se puede partir en dos, una que consiste únicamente de s_n y se denota como 1 en cada caso, y otra para las matrices de los bloques diagonales de cada matriz de 4×4 que corresponden a la base tridimensional (s_a,s_b,s_c). La separación de la representación en conjuntos de matrices de dimensión inferior es definida como *reducción de la representación*, cuya suma directa en este caso es: $\mathbf{D}^{(4)} = \mathbf{D}^{(3)} \oplus \mathbf{D}^{(1)}$. Si $\mathbf{D}^{(3)}$ aun es reducible por una transformación de semejanza, la suma directa corresponde a la reducción

$$\mathbf{D}^{(4)} = \mathbf{D}^{(1)} \oplus \mathbf{D}^{(1)} \oplus \mathbf{D}^{(2)}$$

ya que la reducción de la representación tridimensional es: $\mathbf{D}^{(3)} = \mathbf{D}^{(1)} \oplus \mathbf{D}^{(2)}$. Si $\mathbf{D}^{(2)}$ ya no se puede reducir por una transformación de semejanza, entonces $\mathbf{D}^{(2)}$ es una *representación irreducible*. La representación infiel $\mathbf{D}^{(1)}$ es otra representación irreducible. Por lo tanto, existen tres representaciones irreducibles para el grupo C_{3v}: dos $\mathbf{D}^{(1)}$ y una $\mathbf{D}^{(2)}$. Las letras A y B se utilizan para las *especies de simetría* de representaciones irreducibles unidimensionales, E se emplea para representaciones irreducibles bidimensionales y T se usa para las tridimensionales. La representación infiel unidimensional tiene la lista de caracteres (1,1,1,1,1,1) y pertenece a las especies de simetría definidas como A_1. Las representaciones irreducibles bidimensionales tienen caracteres (2,-1,-1,0,0,0) y su etiqueta es E. La lista completa de caracteres de todas las representaciones irreducibles posibles de un grupo es llamada *tabla de caracteres*. Dado todo esto, para el grupo C_{3v} ya se cuenta con las especies

de simetría A_1 y E. Para construir la especie de simetría faltante se recurre a los teoremas de ortogonalidad. La tabla es la siguiente:

C_{3v}	E	$2C_3$	$3\sigma_v$	h=6	
A_1	1	1	1	z,	x^2+y^2,z^2
A_2	1	1	-1	R_z	
E	2	-1	0	$(x,y)(R_x,R_y)$	(x^2-y^2,xy) (xz,yz)

Por ejemplo, la simetría de una molécula de agua es muy diferente a la de una molécula de metano. Al menos una conformación de la molécula de agua interaccionando con un átomo de metal se puede iniciar en simetría C_{2v} para cálculos no relajados, mientras que al tratar de insertar un átomo de galio en uno de los enlaces CH del metano se ha usado simetría C_s. Los cálculos en moléculas más complejas usualmente recaen en simetría C_1, sin embargo, se debe conocer bien la molécula en estudio para proponer una simetría adecuada para utilizar la simetría correcta. La tabla de caracteres obtenida para un grupo de simetría por medio de alguna base es la que nos permite construir la configuración electrónica del sistema molecular en estudio.

Fuerzas intermoleculares

Introducción

Las fuerzas intermoleculares son la causa de la mayoría de las propiedades físicas y químicas de la materia. Se puede obtener conocimiento cualitativo de dichas fuerzas mediante observaciones generales. El hecho de que los gases se condensen en líquidos sugiere que las fuerzas entre las moléculas deben ser atractivas a grandes separaciones. Por otra parte, el que los líquidos sean tan poco compresibles indica que a pequeñas separaciones de las moléculas, las fuerzas entre ellas deben ser repulsivas.

Se sabe que existen cuatro fuerzas distintas en la naturaleza (Israelachvili, 1992). Dos de estas fuerzas son las interacciones fuertes y las débiles que actúan entre neutrones, protones, electrones y otras partículas elementales. Estas dos fuerzas son de corto alcance, su rango de acción es tan corto que es menor que 10^{-5} nm, y pertenecen al dominio de física nuclear y de altas energías. Las otras dos fuerzas son las interacciones electromagnética y gravitacional, que actúan entre átomos y moléculas (así como entre partículas elementales). Estas fuerzas

de largo alcance son efectivas sobre un intervalo de distancias mucho más grande, de subatómicas a prácticamente distancias infinitas, y consecuentemente son las fuerzas que gobiernan la conducta de las cosas de cada día. Por ejemplo, las fuerzas electromagnéticas —fuente de todas las interacciones intermoleculares— determinan: *i*) las propiedades de sólidos, líquidos y gases, *ii*) la conducta de partículas en solución, *iii*) reacciones químicas y *iv*) la organización de estructuras biológicas. Las fuerzas gravitacionales toman en cuenta el movimiento de las mareas y muchos fenómenos cosmológicos, y cuando actúan junto con las fuerzas intermoleculares determinan fenómenos tales como la altura a la que un líquido se elevará en capilares pequeños y el máximo tamaño que pueden tener animales y árboles.

Fuerza intermolecular y energía potencial de interacción W(r) se corresponden, ya que la fuerza es igual a menos el gradiente del potencial con respecto a la distancia.

Las fuerzas intermoleculares dependen básicamente tanto de los parámetros σ y E_0 como de la distancia *r*, donde σ es el diámetro aparente de la partícula y E_0 es el tamaño de la interacción. Los dos tipos de fuerzas más generales que existen en la naturaleza son las de largo alcance y las de corto alcance. *Las fuerzas de corto alcance* ($r<\sigma$), también conocidas como *fuerzas químicas,* aparecen cuando sus moléculas están tan próximas que sus respectivas nubes electrónicas solapan en una gran extensión. Son de naturaleza repulsiva. La mayoría de los datos de este tipo de fuerzas se obtienen mediante cálculos mecánico-cuánticos. Las *fuerzas de largo alcance* ($r>\sigma$) aparecen cuando las moléculas se hallan a distancias en que no hay solapamiento de sus nubes electrónicas; son de tipo atractivo, y de ellas se tiene un conocimiento más preciso que el de las anteriores. Suelen variar en función inversa de una potencia de la distancia y son de origen puramente electromagnético. Estas fuerzas de largo alcance se dividen en tres tipos: *electrostáticas o de Coulomb, de inducción y de dispersión.* Las fuerzas *electrostáticas* surgen debido a la energía coulombiana asociada con cargas, momentos dipolares, cuadripolares y de orden superior que caracterizan la distribución de la carga en cada molécula. Las fuerzas de *inducción* se deben a la polarización de una molécula neutra por otra molécula cargada o polar. Las fuerzas de *dispersión* son las más importantes de las tres, ya que aparecen incluso en moléculas no polares; su explicación es mecánico-cuántica. London estudió estas últimas y las bautizó con el nombre de fuerzas de dispersión, ya que las oscilaciones que las producen son también las que dan lugar a la dispersión de la luz por las moléculas.

Es importante reconocer que los cambios de estado de sólido a líquido y de líquido a gas son físicos. Ninguna ligadura química se rompe en estos procesos. Cualquiera de los estados hielo, agua o vapor contienen moléculas de H_2O. Cuando el agua hierve para formar vapor, las moléculas individuales permanecen intactas.

Las fuerzas de ligadura que unen los átomos de una molécula se llaman *fuerzas intramoleculares*. Las fuerzas que ocurren entre moléculas, que las pueden condensar para formar un sólido o un líquido, se denominan *fuerzas intermoleculares*.

Típicamente las *fuerzas intermoleculares* son mucho más débiles que las *fuerzas intramoleculares.*

Fuerzas de Coulomb

Las fuerzas intermoleculares que son de un origen puramente electrostático provienen de la fuerza de Coulomb entre las cargas. Las interacciones entre cargas permanentes, dipolos permanentes, cuadrupolos permanentes, etc., que caracterizan la distribución de la carga en cada molécula, caen en esta categoría.

La fuerza de Coulomb entre dos cargas Q y Q_1 separadas por una distancia r es dada por

$$F = -\frac{dW(r)}{dr} = \frac{Q_1 Q}{4\pi\varepsilon\varepsilon_0 r^2}$$

donde ε es la permitividad relativa o constante dieléctrica del medio. Como el campo eléctrico en la posición de Q_1 está dado por $E = \frac{F}{Q_1}$, entonces

$$E = \frac{Q}{4\pi\varepsilon\varepsilon_0 r^2} \qquad (30)$$

Interacción ion-ion

La energía para la interacción de Coulomb entre dos cargas Q_1 y Q está dada por

$$W(r) = \frac{Q_1 Q}{4\pi\varepsilon\varepsilon_0 r} \quad = \quad \frac{z_1 z_2 e^2}{4\pi\varepsilon\varepsilon_0 r}$$

en donde r es la distancia entre las dos cargas Q_1 y Q. La expresión del lado derecho comúnmente se usa para interacciones iónicas, donde la magnitud y el signo de cada carga

iónica está dada en términos de la carga elemental ($e = 1.602x10^{-19}\ C$) multiplicada por la valencia iónica z. Por ejemplo: z = +1 para cationes monovalentes tales como N_a^+, z = –1 para aniones monovalentes tales como Cl^-, z = +2 para cationes divalentes tales como C_a^{2+}, etcétera.

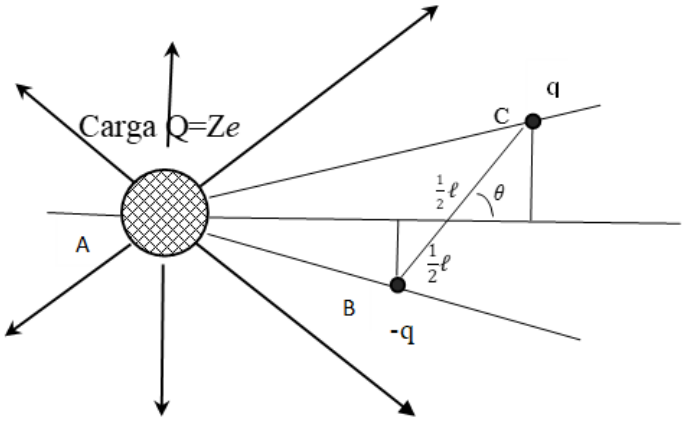
Para cargas del mismo signo, tanto $W(r)$ como $F(r)$ son positivas y la fuerza es repulsiva, mientras que para cargas opuestas tanto la energía $W(r)$ como la fuerza $F(r)$ son negativas y la fuerza es atractiva.

Interacción ion-dipolo

a) Dipolo fijo

Por ejemplo, la interacción electrostática entre un ion de sodio N_a^+ y una molécula polar H_2O. Para derivar el potencial electrostático, la figura 21 muestra una carga Q a una distancia r del centro de una molécula polar de momento dipolar p haciendo un ángulo θ con la línea que une al ion con la molécula.

Figura 21. *Momento dipolar* $p = q\ell\ E$



Fuente: elaboración propia.

Si la longitud del dipolo es ℓ con cargas ±q en cada extremo, entonces la energía de interacción total es la suma de las energías de Coulomb de Q con –q en B y Q con +q en C.

$$W(r,\theta) = -\frac{Qq}{4\pi\varepsilon\varepsilon_0}\left[\frac{1}{AB} - \frac{1}{AC}\right]$$

donde

$$AB = \left[\left(r - \tfrac{1}{2}\ell\cos\theta\right)^2 + \left(\tfrac{1}{2}\ell sin\theta\right)^2\right]^{1/2}$$

$$AC = \left[\left(r + \tfrac{1}{2}\ell\cos\theta\right)^2 + \left(\tfrac{1}{2}\ell sin\theta\right)^2\right]^{1/2}$$

A separaciones r que exceden la longitud dipolar ℓ y al desarrollar en serie de Taylor la raíz, tales distancias pueden escribirse aproximadamente como

$$AB \approx r - \tfrac{1}{2}\ell\cos\theta$$

$$AC \approx r + \tfrac{1}{2}\ell\cos\theta$$

y la energía de interacción en este límite es, entonces

$$\begin{aligned} W(r,\theta) &= -\frac{Qq}{4\pi\varepsilon\varepsilon_0}\left[\frac{1}{r - \tfrac{1}{2}\ell\cos\theta} - \frac{1}{r + \tfrac{1}{2}\ell\cos\theta}\right] \\ &= -\frac{Qq}{4\pi\varepsilon\varepsilon_0}\left[\frac{\ell\cos\theta}{r^2 - \tfrac{1}{4}\ell^2\cos^2\theta}\right] \\ &= -\frac{Qp\cos\theta}{4\pi\varepsilon\varepsilon_0 r^2} \quad = \quad -\frac{(ze)p\cos\theta}{4\pi\varepsilon\varepsilon_0 r^2} \end{aligned} \tag{31}$$

para ℓ despreciable comparada con *r*. Aquí p es el momento dipolar eléctrico definido como

$$p = q\ell\,..$$

Entonces

$$W(r,\theta) = -pE(r)\cos\theta \quad = \quad -\mathbf{p}\cdot\mathbf{E}$$

b) Dipolo rotante

A separaciones grandes o en un medio con una constante dieléctrica ε grande, cuando la energía de interacción es del orden o menor que la energía térmica kT, los dipolos pueden rotar más o menos libremente. Por ejemplo, moléculas en un líquido o gas efectúan movimientos térmicos tanto traslacionales como rotacionales. Sin embargo, aunque los valores de cosθ, senθ, etcétera, al ser promediados son cero, los potenciales promediados sobre el ángulo no son cero, puesto que siempre hay un factor de Boltzmann que da más peso a aquellas orientaciones que tienen una energía más baja (más negativa). Es decir, si una molécula adopta todas las orientaciones con probabilidades iguales la interacción dipolar promedio es cero. Sin embargo, en un promedio térmico las configuraciones de baja energía ocurren preferencialmente cuando son determinadas por un factor de peso de Boltzmann.

Se puede calcular la componente de la energía de interacción que depende de la orientación con la siguiente relación

$$W(r) = \int W(r,\Omega) e^{-W(r,\Omega)/kT} \frac{d\Omega}{4\pi}$$

donde la integración se efectúa sobre los ángulos polar θ y azimutal ϕ, y $d\Omega = \operatorname{sen}\theta d\theta d\phi$. Cuando son integrados sobre todas las orientaciones el resultado de la integral de $d\Omega$ es

$$\int d\Omega = \int_0^{2\pi} d\phi \int_0^{\pi} \operatorname{sen}\theta d\theta = 4\pi$$

Para efectuar el cálculo conviene tener los valores promediados de algunos ángulos:

$$\left\langle \cos^2\theta \right\rangle = \frac{1}{4\pi}\int_0^{\pi} \cos^2\theta \operatorname{sen}\theta d\theta \int_0^{2\pi} d\phi = \frac{1}{3}$$

$$\left\langle \operatorname{sen}^2\theta \right\rangle = \frac{2}{3}$$

$$\left\langle \operatorname{sen}^2\phi \right\rangle = \left\langle \cos^2\phi \right\rangle = \frac{1}{2}$$

$$\left\langle \operatorname{sen}\theta \right\rangle = \left\langle \cos\theta \right\rangle = \left\langle \operatorname{sen}\theta\cos\theta \right\rangle = 0$$

$$\left\langle \operatorname{sen}\phi \right\rangle = \left\langle \cos\phi \right\rangle = \left\langle \operatorname{sen}\phi\cos\phi \right\rangle = 0$$

Cuando $W(r,\Omega)$ es menor que kT se puede desarrollar la función exponencial en serie de Taylor entonces

$$W(r) = \int \left(W(r,\Omega) - \frac{W(r,\Omega)^2}{kT} + \cdots \right) \frac{d\Omega}{4\pi}$$

usando el resultado de la interacción ion-dipolo fijo obtenida anteriormente, se puede conseguir la energía ángulo-promediada para la interacción ion-dipolo rotante:

$$W(r) = \int \left(-\frac{Qp\cos\theta}{4\pi\varepsilon\varepsilon_0 r^2} - \left(\frac{Qp}{4\pi\varepsilon\varepsilon_0 r^2} \right)^2 \frac{\cos^2\theta}{kT} + \cdots \right) \frac{d\Omega}{4\pi}$$

Al sustituir los valores espacialmente promediados de $\left\langle \cos\theta \right\rangle = 0$ y $\left\langle \cos^2\theta \right\rangle = \frac{1}{3}$ mencionados, se obtiene

$$W(r) \approx -\frac{Q^2 p^2}{3(4\pi\varepsilon\varepsilon_0)^2 kTr^4} \qquad \text{para} \qquad kT > \frac{Qp}{4\pi\varepsilon\varepsilon_0 r^2}$$

el cual siempre es atractivo, una propiedad general de términos de segundo orden.

Si tenemos un sistema termodinámico compuesto por muchos de estos dipolos interactuando con cargas, y dado que los dipolos adquieren un gran número de orientaciones angulares, va a existir una entropía del sistema. En este caso, la energía libre A, que nos proporcionará el trabajo disponible, es la que se debe usar para obtener la fuerza al derivarla con respecto a r.

De la relación fundamental de la termodinámica (ver Anexo A, ecuación (A19) con *F=A*):

$$A = U - TS \qquad \text{con} \qquad S = -\left(\frac{\partial A}{\partial T}\right)$$

se puede obtener

$$\left(\frac{\partial\left(A/T\right)}{\partial T}\right) = -\frac{U}{T^2}$$

dado que la $W(r)$ varía como 1/T, se puede integrar la ecuación anterior usando la circunstancia de que $U = \text{const}/T$, en la siguiente forma

$$\frac{A}{T} = -\int const \times T^{-3} dT = const\,\frac{1}{2T^2} = \frac{U}{2T}$$

de donde $A = \frac{U}{2}$, por lo que la energía libre para la interacción ion–dipolo es

$$A = -\frac{Q^2 p^2}{6\left(4\pi\varepsilon\varepsilon_0\right)kTr^4}$$

a partir de la cual se debe calcular la fuerza para tales casos.

Interacción dipolo-dipolo

a) Dipolos fijos

La energía potencial de un dipolo en cualquier campo cuyo potencial es V es igual al trabajo hecho contra este campo en llevar cada carga separadamente desde el infinito hasta un lugar. Si el potencial del campo donde se coloca +q en P_1 es V_1, y –q en P_2 es V_2, entonces

$$W = q\left(V_1 - V_2\right) = q\overline{P_1P_2}\,\frac{\partial V}{\partial s} = p\,\frac{\partial V}{\partial s}$$

donde s va en la dirección del dipolo y *p* es su momento dipolar. En notación vectorial:

$$W = (\mathbf{p} \cdot \nabla)V \tag{32}$$

Si $\mathbf{p}_1$ y $\mathbf{p}_2$ son los momentos vectoriales de dos dipolos A y B y si $\mathbf{r}$ es el vector dirigido desde A hasta B, entonces el potencial en B ocasionado por A, suponiendo que $\mathbf{p}_1$ está en el origen apuntando en la dirección z, es

$$4\pi\varepsilon\varepsilon_0 V = \frac{p_1 \cos\theta}{r^2} = \frac{\mathbf{p}_1 \cdot \mathbf{r}}{r^3} = -\mathbf{p}_1 \cdot \nabla\left(\frac{1}{r}\right) \tag{33}$$

Sustituyendo esta ecuación en la anterior se obtiene

$$\begin{aligned} 4\pi\varepsilon\varepsilon_0 W &= +\mathbf{p}_2 \cdot \nabla\left(\mathbf{p}_1 \cdot \mathbf{r} r^{-3}\right) \\ &= \mathbf{p}_2 \cdot \left(r^{-3}\nabla(\mathbf{p}_1 \cdot \mathbf{r}) + \mathbf{p}_1 \cdot \mathbf{r}\nabla r^{-3}\right) \\ &= \mathbf{p}_2 \cdot \left\{r^{-3}\left[\mathbf{p}_1 \times \nabla \times \mathbf{r} + \mathbf{r} \times (\nabla \times \mathbf{p}_1) + (\mathbf{r} \cdot \nabla)\mathbf{p}_1 + (\mathbf{p}_1 \cdot \nabla)\mathbf{r}\right] - \mathbf{p}_1 \cdot \mathbf{r}\left(3r^{-4}\nabla r\right)\right\} \\ &= \mathbf{p}_2 \cdot \left\{r^{-3}\mathbf{p}_1 - (\mathbf{p}_1 \cdot \mathbf{r})3\mathbf{r}r^{-5}\right\} \end{aligned}$$

Nótese que: $\nabla \times \mathbf{r} = 0$, $\nabla \times \mathbf{p}_1$ también es cero ya que *p* es considerado como puntual y por tanto constante, al igual que $\nabla \mathbf{p}_1$, y que

$$(\mathbf{p}_1 \cdot \nabla)\mathbf{r} = \left(p_{1x}\frac{\partial}{\partial x} + p_{1y}\frac{\partial}{\partial y} + p_{1z}\frac{\partial}{\partial z}\right)(x, y, z) = \mathbf{p}_1$$

Además $\nabla r = \dfrac{\mathbf{r}}{r}$ es un vector unitario.

Por lo tanto

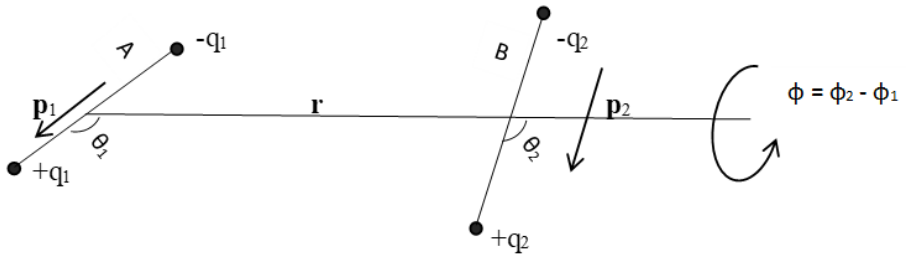
$$4\pi\varepsilon\varepsilon_0 W = r^{-3}\left(\mathbf{p}_2 \cdot \mathbf{p}_1 - 3(\mathbf{p}_1 \cdot \mathbf{r})(\mathbf{p}_2 \cdot \mathbf{r})r^{-2}\right) \tag{34}$$

Si $\mathbf{p}_1$ y $\mathbf{p}_2$ hacen ángulos θ_1 y θ_2 con *r* entre sí, se tiene que

$$W = \frac{p_1 p_2}{4\pi\varepsilon\varepsilon_0 r^3}\left(\cos\varphi - 3\cos\theta_1 \cos\theta_2\right)$$

en donde φ es el ángulo cenital que hace el dipolo $\mathfrak{p}_2$ con el eje z, por lo tanto, respecto al dipolo $\mathfrak{p}_1$. Esto ocurre siempre y cuando ϕ sea el ángulo entre los planos que intersectan en r conteniendo $\mathbf{p}_1$ y $\mathbf{p}_2$ como en la figura 22.

Figura 22. *Trabajo entre dos dipolos fijos*



Fuente: elaboración propia.

Ahora bien, tomando un sistema coordenado tal que **r** está en la dirección de x y $\mathbf{p}_1$ está en el plano *x-y,* los cosenos directores son: $\ell_1 = cos\theta_1,\ \ell_2 = cos\theta_2,\ m_1 = sen\theta_1,$ $m_2 = sen\theta_2 cos\phi,$ y $n_1 = 0$. Entonces

$$\cos\varphi = \ell_1\ell_2 + m_1m_2 + n_1n_2 = \cos\theta_1\cos\theta_2 + \operatorname{sen}\theta_1\operatorname{sen}\theta_2\cos\phi$$

$$\therefore \qquad W = \frac{p_1p_2}{4\pi\varepsilon\varepsilon_0 r^3}\left(\operatorname{sen}\theta_1\operatorname{sen}\theta_2\cos\phi - 2\cos\theta_1\cos\theta_2\right)$$

o bien

$$W = -\frac{p_1p_2}{4\pi\varepsilon\varepsilon_0 r^3}\left(2\cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1\operatorname{sen}\theta_2\cos\phi\right) \qquad (35)$$

b) Dipolo rotante

La energía ángulo-promediada para la interacción dipolo rotante-dipolo rotante se obtiene de la ecuación de promedios de Boltzmann sobre todas las orientaciones $(\theta_1, \theta_2, \phi)$, que en este caso es

$$W(r) = \int\left(W(r,\Omega) - \frac{W(r,\Omega)^2}{kT} + \cdots\right)\frac{d\Omega}{8\pi} \qquad (36)$$

donde $W(r,\Omega) = W(r,\theta_1,\theta_2,\phi)$ está dada por la energía de interacción para el caso dipolo fijo-dipolo fijo obtenida anteriormente en la ecuación (35). Además

$$d\Omega = \int_0^{\pi} \operatorname{sen}\theta_1 d\theta_1 \int_0^{\pi} \operatorname{sen}\theta_2 d\theta_2 \int_0^{2\pi} d\phi = 2\pi\left(-\cos\theta_1\Big|_0^{\pi}\right)\left(-\cos\theta_2\Big|_0^{\pi}\right) = 8\pi$$

Al sustituir la ecuación (35) en la ecuación (36), y definir

$$A(r) = \frac{p_1 p_2}{4\pi\varepsilon\varepsilon_0 r^3}$$

se obtiene

$$W(r) = \int \{-A(r)(2\cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1 \operatorname{sen}\theta_2 \cos\phi) - \frac{A^2(r)}{kT}(2\cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1 \operatorname{sen}\theta_2 \cos\phi)^2\}\frac{d\Omega}{8\pi}$$

como $\int_0^{2\pi} \cos\phi d\phi = \operatorname{sen}\phi\Big|_0^{\pi} = 0$. Entonces dos de las cinco integrales en esta ecuación son cero, y la energía se reduce a

$$W(r) = -\frac{2A(r)}{8\pi}\int_0^{\pi}\cos\theta_1 \operatorname{sen}\theta_1 d\theta_1 \int_0^{\pi}\cos\theta_2 \operatorname{sen}\theta_2 d\theta_2 \int_0^{2\pi} d\phi - \frac{A^2(r)}{8\pi kT}4\int_0^{\pi}\cos^2\theta_1 \operatorname{sen}\theta_1 d\theta_1 \int_0^{\pi}\cos^2\theta_2 \operatorname{sen}\theta_2 d\theta_2 \int_0^{2\pi} d\phi - \frac{A^2(r)}{8\pi kT}\int_0^{\pi}\operatorname{sen}^3\theta_1 d\theta_1 \int_0^{\pi}\operatorname{sen}^3\theta_2 d\theta_2 \int_0^{2\pi}\cos^2\phi d\phi$$

Como

$$\int_0^{\pi}\cos x \operatorname{sen} x dx = \frac{1}{2}\operatorname{sen}^2 x\Big|_0^{\pi} = 0,$$

$$\int_0^{\pi}\cos^2 x \operatorname{sen} x dx = -\frac{1}{3}\cos^3 x\Big|_0^{\pi} = -\frac{1}{3}(-1-1) = \frac{2}{3}$$

$$\int_0^{\pi}\operatorname{sen}^3 x dx = -\cos x\Big|_0^{\pi} + \frac{1}{3}\cos^3 x\Big|_0^{\pi} = -(-1-1) + \frac{1}{3}(-1-1) = \frac{4}{3}$$

así que

$$W(r) = -\frac{A^2(r)}{8\pi kT}4\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)2\pi - \frac{A^2(r)}{8\pi kT}\left(\frac{4}{3}\right)\left(\frac{4}{3}\right)\pi = -\frac{A^2(r)}{kT}\left(\frac{4}{9}+\frac{2}{9}\right)$$

por lo tanto, la energía de interacción es

$$W(r) = -\frac{2p_1^2 p_2^2}{3(4\pi\varepsilon\varepsilon_0)^2 kTr^6}$$

Esta interacción Boltzmann-promediada entre dos dipolos permanentes, conocida como la *interacción de orientación* o *interacción de Keesom*, constituye una de las tres contribuciones importantes a las interacciones de Van der Waals entre átomos y moléculas.

Si se tiene un número muy grande de dipolos rotantes interactuantes, se aplica la deducción correspondiente anterior, y la energía libre de Helmholtz a partir de la cual se puede obtener la fuerza es $A = \frac{U}{2}$, por lo que

$$W(r) = -\frac{p_1^2 p_2^2}{3(4\pi\varepsilon\varepsilon_0)^2 kTr^6}$$

Fuerzas de inducción

El efecto de inducción se debe a la polarización de una molécula neutra por otra molécula cargada o polar. Cuando una molécula neutra se coloca en un campo eléctrico, los electrones son jalados levemente en una dirección mientras que los núcleos son empujados en la dirección opuesta. Este desplazamiento relativo de las cargas en la materia se conoce como *polarización*. Cuando dos moléculas cargadas o dipolares se aproximan entre sí, una polarización mutua toma lugar. Esta polarización de las moléculas puede ser una contribución importante al potencial intermolecular (Hirschfelder et al., 1954).

Considérese un conjunto de cargas ligadas en un campo externo, el cual conduce a un desplazamiento de las cargas. Las diferencias entre los momentos multipolares de la distribución de carga distorsionada y aquella de la distribución sin distorsionar se conocen como los momentos multipolares inducidos. Para campos externos pequeños el momento dipolar inducido puede escribirse de la siguiente manera

$$\mathbf{p}_{ind} = \alpha\mathbf{E} \qquad (37)$$

donde α es la polarizabilidad del objeto. La polarizabilidad es distinta de cero para todos los átomos y moléculas, y juega un papel importante en la determinación de fuerzas coloidales. Para una molécula no-polar la polarizabilidad proviene del desplazamiento de su nube electrónica cargada negativamente relativa a la nuclear cargada positivamente, bajo la

influencia de un campo eléctrico externo. Para moléculas polares hay otras contribuciones a la polarizabilidad, como aquella ocasionada por la rotación de los dipolos permanentes.

La polarización de la materia por la influencia de un campo eléctrico externo se debe a dos efectos: *i*) la polarizabilidad de las moléculas individuales y *ii*) la orientación parcial de momentos dipolares permanentes de las moléculas. El momento dipolar promedio por unidad de volumen debido a estos dos efectos es conocido como la polarización.

Interacción ion-no-polar

Heurísticamente se debe considerar a la energía de interacción entre una carga y un dipolo como la interacción del dipolo ya formado en el campo eléctrico producido por la carga más el trabajo hecho para producir dicho dipolo.

Para encontrar la interacción ion-no-polar, se puede comenzar por modelar la interacción entre un campo eléctrico y una molécula no-polar considerando que esta consiste en dos cargas opuestas ±q interactuando bajo la acción de un resorte de constante k, en analogía a lo que ocurre al aplicar un campo eléctrico a dos cargas opuestas como se ve en la figura 23. En reposo coinciden, pero un campo eléctrico externo generado por un ion adyacente o un dipolo las va a separar una distancia *x.* Así, la energía interna W_{int} para pequeñas deformaciones es

$$W_{\text{int}} = \frac{k}{2}x^2$$

La energía de interacción W_{campo} entre el campo eléctrico **E** y las cargas separadas es

$$W_{campo} = -\mathbf{p}\cdot\mathbf{E} = -qxE \qquad = \qquad -\alpha E^2$$

La energía total de interacción W_{tot} es $W_{\text{int}} + W_{campo} = \left(\frac{k}{2}\right)x^2 - Eqx$

Figura 23. *Interacción ion-no-polar*



Fuente: elaboración propia.

En equilibrio existe un balance entre la interacción externa y la fuerza restauradora del resorte cuando:

$$\left(\frac{\partial W_{tot}}{\partial x}\right) = 0$$

Al derivar se encuentra que en equilibrio la separación de las cargas es

$$x_0 = \frac{Eq}{k}$$

Entonces, dado que el momento dipolar eléctrico es $p = q\ell$ el momento inducido llega a ser

$$p_{ind} = qx_0 = \frac{Eq^2}{k}$$

Al compararlo con la magnitud del momento dipolar $\mathbf{p}_{ind} = \alpha\mathbf{E}$, se demuestra que la polarizabilidad es

$$\alpha = \frac{q^2}{k}$$

Así se ve que conforme la fuerza interna restauradora es más pequeña (i.e., el valor más pequeño de k), la polarizabilidad es más grande. La energía total es

$$W_{tot} = W_{\text{int}} + W_{campo} = \frac{k}{2}\frac{E^2q^2}{k^2} - Eq\frac{Eq}{k}$$

Al simplificar obtenemos

$$W_{tot} = -\frac{1}{2}\alpha E^2$$

Obsérvese que este resultado es independiente de la configuración de cargas que produjo el campo, esto se corrobora en el Anexo B, en donde se deduce esta expresión

directamente pensando en el trabajo que causa el llevar una partícula de un punto a otro. Además, se ve que

$$W_{tot} = \frac{1}{2} W_{campo} = -W_{\text{int}}$$

Para calcular $W_{\text{ion-dipolo inducido}}$ se deben combinar la ecuación (37) con la ecuación (30) para el campo eléctrico asociado con un ion de carga $Q = Ze$ en $\boldsymbol{r} = 0$, esto da como resultado:

$$W_{ion-dipolo\ inducido} = -\frac{\alpha (ze)^2}{2(4\pi\varepsilon\varepsilon_0)^2 r^4}$$

el rango es r^{-4} más bien que r^{-2} ya que un término r^{-2} induce la respuesta, pero un segundo término r^{-2} produce la interacción. Si a este resultado se le agrega la energía libre A de Helmholtz obtenida para la interacción ion-dipolo rotante, se consigue que la polarizabilidad total

$$\alpha_T = \alpha + \frac{p^2}{3kT}$$

Por lo tanto, la energía neta de interacción ion dipolo inducido está dada por

$$W_{neta\ ion-dipolo\ inducido} = -\frac{(ze)^2}{2(4\pi\varepsilon\varepsilon_0)^2 r^4}\left(\alpha + \frac{p^2}{3kT}\right)$$

Interacción dipolo-no-polar (energía de Debye)

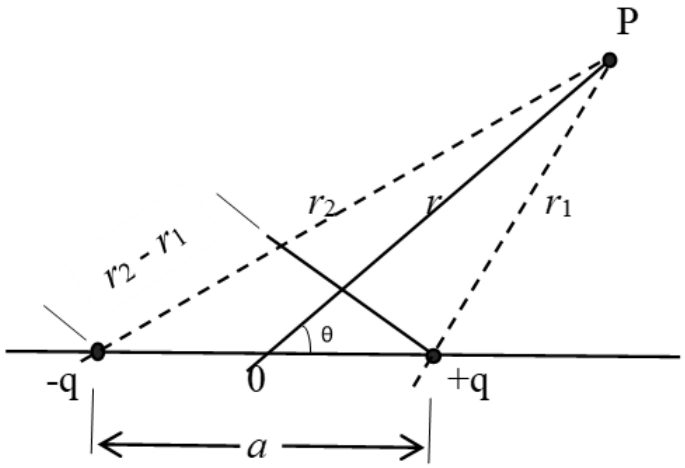
Si se tienen varias cargas $q_1, q_2, q_3, \cdots$, por definición el potencial eléctrico en un punto P es la suma escalar de sus potenciales individuales, esto es

$$V = \frac{1}{4\pi\varepsilon\varepsilon_0}\sum_i \frac{q_i}{r_i}$$

el potencial eléctrico en el punto P debido a un dipolo eléctrico como el de la figura 24 es

$$V = \frac{1}{4\pi\varepsilon\varepsilon_0}\left(\frac{q}{r_1} - \frac{q}{r_2}\right) = \frac{1}{4\pi\varepsilon\varepsilon_0}\frac{q(r_2 - r_1)}{r_1 r_2}$$

Figura 24. *Dipolo eléctrico*



Fuente: elaboración propia.

Si la distancia *a* es muy pequeña comparada con *r,* se puede poner

$$r_1 - r_2 = a\cos\theta \qquad \text{y} \qquad r_1 r_2 = r^2$$

Resultando

$$V = \frac{qa\cos\theta}{4\pi\varepsilon\varepsilon_0 r^2}$$

Como el momento dipolar eléctrico $\boldsymbol{p}$ se define por

$$\boldsymbol{p} = q\boldsymbol{a}$$

donde $\boldsymbol{a}$ es el vector desplazamiento orientado de la carga negativa a la positiva, con lo cual el potencial eléctrico es

$$V = \frac{p\cos\theta}{4\pi\varepsilon\varepsilon_0 r^2}$$

Como el campo eléctrico es menos el gradiente del potencial, en coordenadas polares las componentes del gradiente se pueden expresar como $E_s = -\frac{\partial V}{\partial s}$, $s = x, y, z$, mientras que en coordenadas polares $ds = dr$ para la componente E_r y $ds = rd\theta$ para la componente E_θ, entonces

$$E_r = -\frac{\partial V}{\partial r} = \frac{2p\cos\theta}{4\pi\varepsilon\varepsilon_0 r^3}$$

$$E_r = -\frac{1}{r}\frac{\partial V}{\partial \theta} = \frac{p\,\mathrm{sen}\,\theta}{4\pi\varepsilon\varepsilon_0 r^3}$$

$$\mathbf{E} = E_r \hat{i}_r + E_\theta \hat{i}_\theta = \frac{2p\cos\theta}{4\pi\varepsilon\varepsilon_0 r^3}\hat{i}_r + \frac{p\,\mathrm{sen}\,\theta}{4\pi\varepsilon\varepsilon_0 r^3}\hat{i}_\theta$$

La magnitud del campo eléctrico se calcula como sigue

$$|E|^2 = \frac{4p^2\cos^2\theta}{(4\pi\varepsilon\varepsilon_0)^2 r^6} + \frac{p^2\,\mathrm{sen}^2\,\theta}{(4\pi\varepsilon\varepsilon_0)^2 r^6} = \frac{p^2}{(4\pi\varepsilon\varepsilon_0)^2 r^6}\left(4\cos^2\theta + \mathrm{sen}^2\,\theta\right)$$

$$\therefore \qquad E = \frac{p}{4\pi\varepsilon\varepsilon_0 r^3}\left(3\cos^2\theta + 1\right)^{1/2}$$

Al sustituir este valor en la energía de interacción se encuentra que para un dipolo fijo:

$$W_{dipolo-inducido\ dipolo}(r,\theta) = -\frac{1}{2}\alpha E^2 = -\frac{p^2\alpha\left(1+3\cos^2\theta\right)}{2(4\pi\varepsilon\varepsilon_0)^2 r^6}$$

La interacción conocida como de Debye se obtiene promediando angularmente la relación anterior. Puesto que el promedio angular de $\cos^2\theta$ es 1/3, entonces para un dipolo rotante se obtiene finalmente

$$W(r) = \int W(r,\Omega)\frac{d\Omega}{4\pi} = -\frac{p^2\alpha}{(4\pi\varepsilon\varepsilon_0)^2 r^6} \tag{38}$$

Este resultado representa la interacción atractiva de Debye y constituye la segunda de las tres contribuciones inversas de la sexta potencia a la energía de interacción total de Van der Waals entre moléculas.

Fuerzas de dispersión de London o de Van der Waals

Las fuerzas de dispersión son de naturaleza mecánico-cuántica y fueron estudiadas por London. Debido a que siempre están presentes en átomos y moléculas, las fuerzas de dispersión son más importantes que las anteriores, además juegan un papel en una multitud de fenómenos importantes como adhesión, tensión superficial, adsorción física, mojado, las propiedades de gases, líquidos y películas delgadas, la fuerza de los sólidos, la floculación de partículas en líquidos, y la estructura de macromoléculas condensadas tales como proteínas y polímeros. Sus principales características pueden resumirse como sigue:

- Son fuerzas de largo alcance y, dependiendo de la situación, pueden ser efectivas a grandes distancias (mayores de 10 nm) y hasta distancias de espaciamientos interatómicos (cerca de 0.2 nm).

- Estas fuerzas pueden ser repulsivas o atractivas, y en general las fuerzas de dispersión entre dos moléculas (o partículas grandes) no siguen una ley de potencias simple.
- Las fuerzas de dispersión no solamente juntan moléculas sino también tienden a alinearlas mutuamente u orientarlas, aunque usualmente este efecto orientante es débil.
- La interacción de dispersión de dos cuerpos es afectada por la presencia de otros cuerpos cercanos. Esto es conocido como la no-aditividad de una interacción.

Las fuerzas de dispersión son amenas para una multitud de tratamientos teóricos de diversas complejidades, los más rigurosos de los cuales nos llevarían al mundo de la electrodinámica cuántica. Su origen puede ser entendido intuitivamente como sigue: para un átomo no-polar tal como helio, el tiempo promedio de su momento dipolar es cero, aun en cualquier instante existe un momento dipolar finito dado por las posiciones instantáneas de los electrones alrededor de los protones nucleares. Este dipolo instantáneo genera un campo eléctrico que polariza cualquier átomo neutral cercano, induciendo un momento dipolar en él. La interacción resultante entre los dos dipolos conduce a una fuerza atractiva instantánea entre los dos átomos y el tiempo promedio de esta fuerza es finita.

La ecuación de London

a) Deducción heurística

En el átomo de Bohr un electrón gira alrededor del protón. La distancia más pequeña entre el electrón y el protón se conoce como el primer radio de Bohr a_o y es el radio en el cual la energía de Coulomb $e^2/4\pi\varepsilon_0 a_0$ es igual a $2h\nu$, es decir,

$$a_0 = \frac{e^2}{2(4\pi\varepsilon_0)h\nu} = 0.053nm \qquad (39)$$

donde h es la constante de Planck y ν es la frecuencia orbital del electrón. Para un átomo de Bohr, $\nu = 3.3 \times 10^{15}\ s^{-1}$, así que $h\nu = 2.2 \times 10^{-18}$ J. Esta es la energía de un electrón en el primer radio de Bohr y es igual a la energía para ionizar el átomo —el *primer potencial de ionización,* I—.

El átomo de Bohr no tiene momento dipolar permanente. Sin embargo, en cualquier instante existe un momento dipolar instantáneo

$$p = a_0 e$$

cuyo campo polarizará un átomo neutral cercano produciendo una interacción atractiva que es completamente análoga a la interacción dipolo-inducido dipolo discutida anteriormente. Por lo tanto, la energía de esta interacción en el vacío está dada por la ecuación (38)

$$W(r) = -\frac{p^2 \alpha}{(4\pi\varepsilon\varepsilon_0)^2 r^6} = \frac{(a_0 e)^2 \alpha}{(4\pi\varepsilon\varepsilon_0)^2 r^6}$$

donde α es la polarizabilidad electrónica del segundo átomo de Bohr.

Por otra parte, regresando a la figura 23 de la interacción ion-no polar, la fuerza externa F_{ext} del electrón debida al campo E es

$$F_{ext} = eE$$

la cual debe ser balanceada en equilibrio por la fuerza atractiva entre la órbita del electrón desplazada y el núcleo. Esto no es otra cosa que la fuerza de Coulomb $e^2 / 4\pi\varepsilon\varepsilon_0 r^2$ resuelta a lo largo de la dirección del campo. La fuerza interna (restauradora) es, por lo tanto

$$F_{\text{int}} = \frac{e^2}{4\pi\varepsilon\varepsilon_0 r^2} \operatorname{sen}\theta \approx \frac{e^2 \ell}{4\pi\varepsilon\varepsilon_0 r^3} \approx \frac{e}{4\pi\varepsilon\varepsilon_0 r^3} p_{ind}$$

en equilibrio $F_{ext} = F_{\text{int}}$ lo cual conduce a

$$p_{ind} = 4\pi\varepsilon\varepsilon_0 r^3 E = \alpha E$$

por lo tanto, la polarizabilidad es $\alpha = 4\pi\varepsilon\varepsilon_0 r^3$. Como en este caso el radio corresponde con el radio de Bohr, se tiene que

$$a_0^3 = \frac{\alpha}{4\pi\varepsilon\varepsilon_0}$$

además de la ecuación (39) se obtiene

$$a_0^3 = \frac{e^2 a^2}{2(4\pi\varepsilon\varepsilon_0) h\nu}$$

al igualar estas dos ecuaciones resulta que $\alpha h\nu \approx (ea)^2$, por lo tanto la energía de interacción se puede escribir como

$$W(r) \approx -\frac{\alpha^2 h\nu}{(4\pi\varepsilon\varepsilon_0)^2 r^6}$$

Excepto por un factor numérico, esta ecuación es la misma que aquella que fue derivada por London en 1930 usando teoría de perturbaciones de mecánica cuántica. La famosa expresión de London (1937) para la energía de interacción de dispersión entre dos átomos o moléculas idénticas es

$$W(r) = \frac{C_{disp}}{r^6} = -\frac{3}{4}\frac{\alpha^2 h\nu}{(4\pi\varepsilon\varepsilon_0)^2 r^6} = -\frac{3}{4}\frac{\alpha^2 I}{(4\pi\varepsilon\varepsilon_0)^2 r^6}$$

y para dos átomos diferentes

$$W(r) = -\frac{3}{2}\frac{\alpha_1\alpha_2}{(4\pi\varepsilon\varepsilon_0)^2 r^6}\frac{h\nu_1\nu_2}{(\nu_1+\nu_2)} = -\frac{3}{2}\frac{\alpha_1\alpha_2}{(4\pi\varepsilon\varepsilon_0)^2 r^6}\frac{I_1 I_2}{(I_1+I_2)} \qquad (40)$$

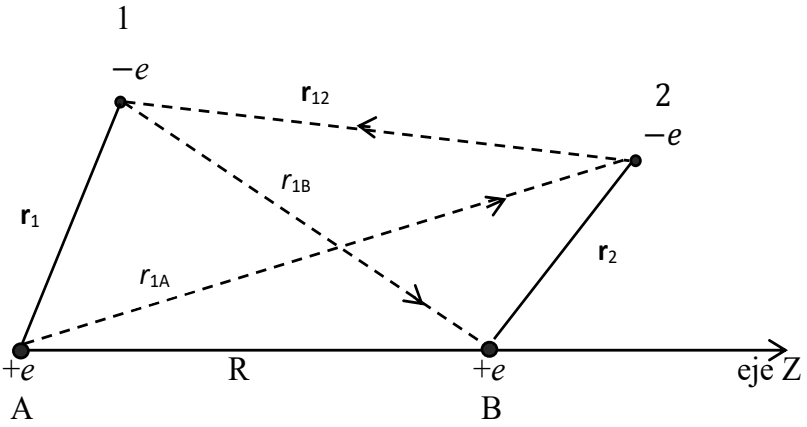
Con este simple modelo se ve que mientras las fuerzas de dispersión son mecánico-cuánticas (al determinar los momentos dipolares instantáneos, pero fluctuantes, de átomos neutros), la interacción resultante aún es electrostática, una clase de fuerza de polarización mecánico-cuántica. Nótese además que la dependencia $\frac{1}{r^6}$ de la distancia es la misma que aquella para las otras dos interacciones de polarización (las fuerzas de Keesom y las de Debye) que contribuyen a la fuerza neta de Van der Waals.

b) Deducción exacta

La teoría de London se ilustra de una manera más sencilla tomando el ejemplo de dos átomos de hidrógeno. La interacción de largo alcance de Van der Waals entre dos átomos de hidrógeno en sus estados base se puede calcular por medio de la teoría de perturbaciones.

Supongamos que los núcleos de los dos átomos de hidrógeno A y B están fijos en el espacio alejados una distancia R, y que el eje Z se elige paralelo a la línea que va de A a B. Entonces si $\mathbf{r}_1$ es el vector de desplazamiento del electrón 1 del núcleo A y $\mathbf{r}_2$ es el vector desplazamiento del electrón 2 de núcleo B como en la figura 25.

Figura 25. *Dipolos de dos átomos de hidrógeno*



Fuente: elaboración propia.

El operador Hamiltoniano del sistema de interacción de dos átomos A y B con un electrón cada uno puede escribirse de la siguiente forma

$$H = H_0 + H'$$

donde

$$H_0 = -\frac{\hbar^2}{2m}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{e^2}{r_1} - \frac{e^2}{r_2}$$

y

$$H' = \frac{e^2}{R} - \frac{e^2}{r_{12}} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}}$$

cuando $r_1, r_2 << R$. Nótese que el operador hamiltoniano *H* para la interacción entre dos átomos de hidrógeno es el mismo que para la interacción entre dos dipolos eléctricos formados de un electrón y un núcleo cada uno de ellos. La ecuación (34) es la de un dipolo eléctrico fijo.

El hamiltoniano no perturbado H_0 tiene como solución la función de onda (ver Anexo C):

$$u_0(\mathbf{r}_1, \mathbf{r}_2) = u_{100}(\mathbf{r}_1) u_{100}(\mathbf{r}_2) = \frac{1}{\pi a_0^3} e^{-(r_1 + r_2)/a_0}$$

para dos átomos de hidrógeno no-interactuantes en sus estados base.

Los términos de interacción H' pueden ser tratados como una perturbación. De manera que los efectos de H' en cualquiera de los estados no-perturbados de los dos átomos

pueden estudiarse por teoría de perturbaciones independientes del tiempo. Si ambos átomos están en el estado base, el cambio en energía debido a la perturbación es

$$\Delta E = \langle 0|H'|0\rangle + \sum_n \frac{|\langle 0|H'|n\rangle|^2}{E_0 - E_n}$$

con

$$\Delta E = W - W_0, \;\; W_0 = E_m, \;\; W_1 = H'_{mm} = \langle m|H'|m\rangle, \;\; W_2 = \sum_n{}' \frac{|H'_{mn}|^2}{E_m - E_n} = \sum_n{}' \frac{|\langle m|H'|m\rangle|^2}{E_m - E_n}$$

y, m = 0 (ver Anexo C). Si R es grande comparado con el radio de Bohr a_o, la mayor contribución a las integrales en los elementos de matriz $\langle 0|H'|0\rangle$ y $\langle 0|H'|n\rangle$ vendrá de la vecindad $r_j \approx a_0, \;\; j = 1,2$. En consecuencia H' puede expresarse como un desarrollo multipolar en potencias de 1/R, escogiendo el eje z a lo largo de R, y manteniendo los términos más bajos. Para desarrollar en serie de Taylor, entonces, se hace la siguiente identificación de vectores:

$$\boldsymbol{r}_{1B} = \boldsymbol{R} - \boldsymbol{r}_1 \qquad \boldsymbol{r}_{1A} = \boldsymbol{R} + \boldsymbol{r}_2$$
$$\boldsymbol{r}_{12} = \boldsymbol{r}_1 - \boldsymbol{r}_{1A} = \boldsymbol{r}_1 - \boldsymbol{R} - \boldsymbol{r}_2 = -\boldsymbol{R} - (\boldsymbol{r}_2 - \boldsymbol{r}_1)$$

u

$$r_{1A} = |\boldsymbol{R} - \boldsymbol{r}_2| = \sqrt{R^2 + r_2^2 - 2\boldsymbol{R}\cdot\boldsymbol{r}_2}$$
$$= R\left(1 - \frac{2z_2}{R} + \frac{r_2^2}{R^2}\right)^{1/2}$$

$$r_{12} = |-\boldsymbol{R} - (\boldsymbol{r}_2 - \boldsymbol{r}_1)| = \sqrt{R^2 + |\boldsymbol{r}_2 - \boldsymbol{r}_1|^2 - 2\boldsymbol{R}\cdot(\boldsymbol{r}_2 - \boldsymbol{r}_1)}$$
$$= R\left(1 + \frac{2(z_2 - z_1)}{R} + \frac{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^2}{R^2}\right)^{1/2}$$

es decir, el hamiltoniano de la perturbación está dado por

$$H' = \frac{e^2}{R}\left\{1 + \left[1 + \frac{2(z_2 - z_1)}{R} + \frac{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}{R^2}\right]^{-1/2} - \left(1 - \frac{2z_1}{R} + \frac{r_1^2}{R^2}\right)^{-1/2} - \left(1 + \frac{2z_2}{R} + \frac{r_2^2}{R^2}\right)^{-1/2}\right\}$$

Al desarrollar en serie de Taylor usando $(1 \pm x)^{-\frac{1}{2}} = 1 \mp \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 \cdots$, se obtiene

$$H' \cong \frac{e^2}{R}\left\{1+1-\frac{2(z_2-z_1)}{2R}-\frac{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}{2R^2}\right.$$
$$+\frac{3}{8}\left(\frac{2(z_2-z_1)}{R}+\frac{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}{R^2}\right)^2-1-\frac{z_1}{R}+\frac{r_1^2}{2R^2}$$
$$\left.-\frac{3}{8}\left(\frac{2z_1}{R}+\frac{r_1^2}{R^2}\right)^2-1+\frac{z_2}{R}+\frac{r_2^2}{2R^2}-\frac{3}{8}\left(\frac{2z_2}{R}+\frac{r_2^2}{R^2}\right)^2+\cdots\right\}$$

$$H' \cong \frac{e^2}{R}\left\{-\frac{x_1^2+y_1^2+z_1^2+x_2^2+y_2^2+z_2^2-2x_1x_2-2y_1y_2-2z_1z_2}{2R^2}+\frac{3}{8}\left[4\left(z_2^2+z_1^2-2z_1z_2\right)+O^s\right]\right.$$
$$\left.+\frac{r_1^2}{2R^2}-\frac{3}{8}\left(\frac{4z_1^2}{R^2}+\frac{r_1^4}{R^4}+\frac{4z_1r_1^2}{R^3}\right)+\frac{r_2^2}{2R^2}-\frac{3}{8}\left(\frac{4z_2^2}{R^2}+\frac{r_2^4}{R^4}+\frac{4z_2r_2^2}{R^3}\right)+\cdots\right\}$$

$$H' \cong \frac{e^2}{R^3}\left(x_1x_2+y_1y_2+z_1z_2-3z_1z_2\right)$$

$$H' \cong \frac{e^2}{R^3}\left(x_1x_2+y_2y_1-2z_1z_2\right) \qquad (41)$$

por lo tanto, la perturbación H' es la energía de interacción de dos dipolos eléctricos que corresponden a las configuraciones instantáneas de los dos átomos.

Es evidente a la vez que el valor esperado del término principal en H' para el estado $u_0(\mathbf{r}_1,\mathbf{r}_2)$ es cero, ya que u_0 es una función par de $\mathbf{r}_1$ y $\mathbf{r}_2$, y H' es una función impar de $\mathbf{r}_1$ y $\mathbf{r}_2$ separadamente. También puede demostrarse que todos los términos superiores despreciados en H' tienen valor esperado cero para u_0, puesto que estos términos pueden ser expresados como armónicos esféricos de orden diferente de cero. Entonces el término principal de la energía de interacción es la perturbación a segundo-orden del término dipolo-dipolo, que es proporcional a H'^2 y, en consecuencia, varía como $\frac{1}{R^6}$.

La corrección a primer orden desaparece[1] (Saxon, 1968) cuando se utiliza la ecuación (41) ya que u_o es una función *par* de $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$, mientras que H' es una función *impar* de $\boldsymbol{r}_1$ y $\boldsymbol{r}_2$. El término de segundo orden da

[1] Como el hamiltoniano no perturbado H_o es simétrico, y el término perturbativo H' tiene simetría definida, *par o impar*, se concluye que la mitad de los elementos de matriz H' son cero. Si H' es *par*, todos los elementos de matriz entre estados de paridad opuesta se anulan automáticamente:

$$\left\langle u_{par}\left|H'_{par}\right|u_{impar}\right\rangle=\left\langle u_{impar}\left|H'_{par}\right|u_{par}\right\rangle=0$$

$$\Delta E = \frac{e^4}{R^6}\sum_{m,n}\frac{\left|x_{0,m}x_{o,n} + y_{0,m}y_{o,n} - 2z_{0,m}z_{o,n}\right|^2}{2E_0 - E_m - E_n} \tag{42}$$

Aquí E_0 y E_n son las energías del estado base y del n-ésimo estado excitado del átomo de hidrógeno, y los elementos de matriz $x_{0,n}$, $y_{0,n}$ y $z_{0,n}$ se toman entre estos estados (Mahanty y Ninham, 1976). La ecuación (42) indica que ΔE está acotado por dos valores extremos, y su rango de valores debe estar entonces entre

$$(\Delta E)_{\text{inf}} \le \Delta E \le (\Delta E)_{\text{sup}}.$$

Se pueden obtener los valores $(\Delta E)_{\text{inf}}$ y $(\Delta E)_{\text{sup}}$ por los métodos perturbativo y variacional, respectivamente (Schiff, 1955). Nótese que ΔE puede ser identificada con la energía de interacción $W(R)$.

Cálculo perturbativo

De la ecuación B12 del Anexo B, el cambio a segundo orden en la energía de los dos átomos de hidrógeno es

$$W(R) = \sum{}' \frac{\left|H'_{0n}\right|^2}{E_0 - E_n} \tag{43}$$

donde el índice *n* se refiere a todos los estados del par de átomos de hidrógeno no-perturbados (incluyendo estados disociados), y el estado base u_0 es excluido de la sumatoria e integración que es denotada por $\sum'$. Es evidente que $W(R)$ es negativo, puesto que $E_0 < E_n$ y el numerador de cada término en (43) es positivo. Entonces, se concluye que la interacción es atractiva y proporcional a $\frac{1}{R^6}$ cuando *R* es grande; se puede demostrar que

---

Análogamente, si *H´* es *impar*, todos los elementos de matriz entre estados de la misma paridad se anulan automáticamente:

$$\left\langle u_{par}\left|H'_{impar}\right|u_{par}\right\rangle = \left\langle u_{impar}\left|H'_{impar}\right|u_{impar}\right\rangle = 0$$

Estas reglas tan generales son casos especiales de las llamadas reglas de selección en espectroscopía. Hay que observar que las reglas de selección para *H´* impar tienen como consecuencia que la corrección a primer orden para la energía se anula.

estas dos conclusiones son válidas para cualquier par de átomos que están en estados base esféricamente simétricos no-degenerados.

Se puede obtener un límite superior de la cantidad positiva $-W(R)$ reemplazando cada E_n en (43) por la energía E_{n^*} del estado excitado más bajo de los dos átomos de hidrógeno para los cuales H'_{0n^*} es diferente de cero. Entonces el denominador puede ser sacado de la sumatoria, la cual se puede evaluar como un producto de matrices

$$\sum{}' \left|H'_{0n}\right|^2 = \sum{}' H'_{0n}H'_{0n} = \sum H'_{0n}H'_{0n} - H'^2_{00} = \left(H'^2\right)_{00} - H'^2_{00} \tag{44}$$

donde la prima en la sumatoria denota omisión del término $n = 0$. Puesto que se ha visto que $H'_{00} = 0$, se tiene que

$$-W(R) \le \frac{\left(H'^2\right)_{00}}{E_{n^*} - E_0} \tag{45}$$

el estado n^* es aquel en el cual ambos átomos están excitados a estados del número cuántico principal 2, así que (Anexo D):

$$E_0 = -2\left(e^2/a_0\right), \quad E_{n^*} = -2\left(e^2/8a_0\right), \quad y \quad E_{n^*} - E_0 = 3e^2/4a_0$$

De la ecuación (41) se tiene que

$$H'^2 = \frac{e^2}{R^6}\left(x_1^2x_2^2 + y_1^2y_2^2 + 4z_1^2z_1^2 + 2x_1x_2y_1y_2 - \cdots\right)$$

El valor esperado de los términos cruzados como $x_1x_2y_1y_2$ es cero puesto que estos términos son funciones impares de una de las componentes cartesianas de $\mathbf{r}_1$ o $\mathbf{r}_2$. Cada uno de los primeros tres términos en el paréntesis de la ecuación anterior son el producto de dos factores idénticos que son iguales a

$$\int x^2\left|u_{100}(\mathbf{r})\right|^2 d\tau = \frac{1}{3}\int r^2\left|u_{100}(\mathbf{r})\right|^2 d\tau$$

$$= \frac{1}{3\pi a_0^3}\int_0^\infty r^2 e^{-\frac{2r}{a_0}} 4\pi r^2 dr$$

esta integral tiene solución en términos de funciones gama, $\int_0^\infty t^{z-1}\, e^{-kt} dt = \frac{\Gamma(z)}{k^z}$, donde $\Gamma(z+1) = 1 \cdot 2 \cdot 3 \cdots (n-1)n = n!$, entonces

$$\frac{1}{3\pi a_0^3}\int_0^\infty r^2 e^{-\frac{2r}{a_0}} 4\pi r^2 dr = \frac{4}{3a_0^3}\left(\frac{a_0}{2}\right)^5 \Gamma(5)$$
$$= a_0^2$$

así que $\left(H'^2\right)_{00} = \frac{6e^4 a_0^4}{R^6}$. Al hacer la sustitución en la ecuación (45) se tiene $W(R) \geq -\frac{8e^2 a_0^5}{R^6}$

Por lo tanto $(\Delta E)_{\inf} = -\frac{8e^2 a_0^5}{R^6}$.

Cálculo Variacional

Para obtener el límite superior de $W(R)$ por el método variacional se debe elegir una función de prueba ψ. Si tal función no depende de R, la dependencia de R del valor esperado de la energía será como aquella de H', esto es, $1/R^6$. Un límite superior con esta dependencia de R no es de ningún valor para este caso, puesto que lo que realmente se quiere determinar es un límite del coeficiente de la interacción $1/R^6$. Una elección útil para ψ es aquella en la cual hay un término proporcional a H', ya que entonces habrá términos en el valor esperado que son proporcionales a H'^2 y en consecuencia variará como $1/R^6$. Como función de prueba se escoge (Anexo D):

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = u_{100}(\mathbf{r}_1) u_{100}(\mathbf{r}_2)(1 + AH')$$

donde A es el parámetro de variación. Para normalizar la función de prueba se usa la siguiente condición de normalización

$$E_0 \leq \frac{\int \psi^* H \psi d\tau}{\int |\psi|^2 d\tau}$$

entonces se obtiene

$$E_0 + W(R) \leq \frac{\iint u_0(1 + AH')\left(H_0 + H'\right)u_0(1 + AH')d\tau_1 d\tau_2}{\iint u_0^2(1 + AH')^2 d\tau_1 d\tau_2} \qquad (46)$$

donde u_0 es el producto de los estados base de las funciones de onda del átomo de hidrógeno, y A es real. Después de integrar, el lado derecho de la ecuación (46) se reduce a

$$\frac{E_0 + 2A\left(H'^2\right)_{00} + A^2\left(H'H_0H'\right)_{00}}{1 + A^2\left(H'^2\right)_{00}} \tag{47}$$

puesto que u_0 es una eigenfunción normalizada de H_0 con eigenvalor

$$E_0 = -\frac{e^2}{a_0}$$

y $H'_{00} = \left(H'^3\right)_{00} = 0.$ Fácilmente se ve que $\left(H'H_0H'\right)_{00}$ es una suma de cuadrados de factores de la forma $\int u_{100}(\mathbf{r}) x H_0 x u_{100}(\mathbf{r}) d\tau$; por cómputo directo se puede demostrar que esto es cero.

Puesto que se está interesado solo en términos de orden H'^2, se expande el denominador de la ecuación (47):

$$\left[E_0 + 2A\left(H'^2\right)_{00}\right]\left[1 + A^2\left(H'^2\right)_{00}\right]^{-1} \cong E_0 + \left(H'^2\right)_{00}\left(2A - E_0A^2\right)$$

Observando que E_0 es negativa, se encuentra que esta ecuación tiene un mínimo con respecto a la variación de A cuando $A = 1/E_0$ en este caso la ecuación (46) se convierte en

$$E_0 + W(R) \le E_0 + \frac{\left(H'^2\right)_{00}}{E_0} = E_0 - \frac{6e^2a_0^5}{R^6}$$

Por lo tanto $(\Delta E)_{\text{sup}} = -\frac{6e^2a_0^5}{R^6}$. Finalmente, el rango de energías de interacción está entre

$$-\frac{8e^2a_0^5}{R^6} \le \Delta E \le -\frac{6e^2a_0^5}{R^6}$$

Una estimación cuidadosa de la sumatoria en la ecuación (42) da (Pauling y Beach, 1935)

$$W(R) = -6.5\frac{e^2a^5}{R^6} \tag{48}$$

London (1930) reconoció que el lado derecho de la ecuación (42) puede expresarse en términos de los valores *f* de dispersión o fuerza del oscilador. Para una transición dipolar entre dos estados $|\ell\rangle$ y $|m\rangle$, la fuerza del oscilador está definida como

$$f_{\ell m} = \frac{2m}{\hbar^2}\left(E_m - E_\ell\right)\left|z_{\ell m}\right|^2 \tag{49}$$

Al tratar con productos del tipo $x_{0,m}x_{0,n}y_{0,m}y_{0,n}$ siempre es posible escoger *m* y *n* tales que solamente uno de los elementos $x_{0,m}$ $y_{0,m}$ y $z_{0,m}$ no desaparecen. De esta manera

$$W(R)=\frac{e^4}{R^6}\sum_{m,n}\frac{x_{0,m}^2x_{0,n}^2+y_{0,m}^2y_{0,n}^2+4z_{0,m}^2z_{0,n}^2}{\left(E_0-E_m\right)+\left(E_0-E_n\right)}$$

por consideraciones de simetría, se obtiene

$$W(R)=\frac{6e^4}{R^6}\sum_{m,n}\frac{z_{0,m}^2z_{0,n}^2}{\left(E_0-E_m\right)+\left(E_0-E_n\right)} \tag{50}$$

y al utilizar la ecuación (49) se tiene

$$\begin{aligned}W(R)&=\frac{3}{2}\frac{\hbar^4e^4}{R^6m^2}\sum_{m,n}\frac{f_{0m}f_{0n}}{\left(E_m-E_0\right)\left(E_n-E_0\right)}\frac{1}{\left(E_0-E_m\right)+\left(E_0-E_m\right)}\\&=\frac{3}{2}\frac{\hbar e^4}{R^6m^2}\sum_{m,n}\frac{\hbar^2f_{0m}f_{0n}}{\left(E_m-E_0\right)\left(E_n-E_0\right)}\frac{1}{\frac{1}{\hbar}\left[\left(E_0-E_m\right)+\left(E_0-E_n\right)\right]}\\&=-\frac{3}{2}\frac{\hbar e^4}{R^6m^2}\sum_{m,n}\frac{f_{0m}f_{0n}}{\frac{\left(E_m-E_0\right)}{\hbar}\frac{\left(E_n-E_0\right)}{\hbar}}\frac{1}{\frac{\left(E_m-E_0\right)}{\hbar}+\frac{\left(E_n-E_0\right)}{\hbar}}\end{aligned}$$

Como

$$\omega_m = \frac{E_m - E_0}{\hbar} \quad \text{y} \quad \omega_n = \frac{E_n - E_0}{\hbar}$$

corresponden a las frecuencias de la radiación electromagnética que causaría la transición del estado base $|0\rangle$ al estado excitado $|m\rangle$ en la molécula aislada (Hunter, 1987), entonces

$$W(R) = -\frac{3\hbar e^4}{2R^6 m^2} \sum_{m,n} \frac{f_{0m} f_{0n}}{\omega_m \omega_n (\omega_m + \omega_n)}$$

Esta ecuación, aunque fue derivada para el ejemplo del átomo de hidrógeno, es de validez general. Para un átomo o molécula con *n* electrones f_{lm} satisface la regla de la suma (Landau y Lifshitz, 1958):

$$\sum_m f_{\ell m} = n$$

$f_{\ell m}$ se puede obtener de datos experimentales de intensidad de líneas espectrales (Garstang, 1966). Tomando $\sum_j f_{0j} = n_j$ se tiene que

$$W(R) = -\frac{3\hbar}{2R^6} \frac{e^4}{m^2} \frac{n_m n_n}{\omega_m \omega_n (\omega_m + \omega_n)} \qquad (51)$$

Si tomamos dos moléculas con n_1 y n_2 electrones que tienen transiciones ω_1 y ω_2 respectivamente, la ecuación (50) finalmente se reduce a la forma

$$W(R) \cong -\frac{3e^4\hbar}{2m^2 R^6} \frac{n_1 n_2}{\omega_1 \omega_2 (\omega_1 + \omega_2)}$$

La forma de *W*(*R*) en la ecuación (51), originalmente dada por London, es particularmente útil en la estimación de la fuerza de dispersión entre moléculas de datos espectroscópicos experimentales. Cambiando los subíndices *m* por *i*, y *n* por *j*, la ecuación (51) se puede volver a escribir como

$$W(R) = -\frac{3\hbar}{2R^6} \frac{e^2 n_i}{m} \frac{e^2 n_j}{m} \frac{1}{\omega_i \omega_j (\omega_i + \omega_j)} \qquad (52)$$

al usar la identidad

$$\frac{1}{ab(a+b)} = \frac{2}{\pi} \int_0^\infty \frac{d\xi}{(a^2 + \xi^2)(b^2 + \xi^2)}$$

se encuentra que

$$W(R) = -\frac{3\hbar}{2R^6}\frac{e^2 n_i}{m}\frac{e^2 n_j}{m}\frac{2}{\pi}\int_0^\infty \frac{d\xi}{\left(\omega_i^2+\xi^2\right)\left(\omega_j^2+\xi^2\right)}$$

Definiendo $\omega = i\xi \Rightarrow d\omega = id\xi$

$$\omega^2 = i^2\xi = -\xi^2$$

se obtiene

$$W(R) = -\frac{3\hbar}{2\pi R^6}\frac{e^2 n_i}{m}\frac{e^2 n_j}{m}2\int_0^\infty \frac{d\,\omega/i}{\left(\omega_i^2+\omega^2\right)\left(\omega_j^2+\omega^2\right)}$$

$$= -\frac{3\hbar}{\pi R^6}\int_0^\infty \frac{d\omega}{i}\frac{e^2 n_i}{m\left(\omega_i^2+\omega^2\right)}\frac{e^2 n_j}{m\left(\omega_j^2+\omega^2\right)}$$

De acuerdo con la teoría clásica de campos, las ecuaciones de Maxwell se usan para definir la polarizabilidad dinámica (Blokhintsev, 1964; Jackson, 1962):

$$\alpha(\omega) = \frac{e^2}{m}\sum_n \frac{f_{0n}}{\omega_{n0}^2-\omega^2}$$

Si uno de los osciladores predomina sobre los demás, esta ecuación se puede escribir como

$$\alpha_j(\omega) \cong \frac{e^2 n_j}{m\left(\omega_j^2-\omega^2\right)}, \qquad j = 1, 2 \qquad (53)$$

entonces $$W(R) = -\frac{3\hbar}{\pi R^6}\int_0^\infty \frac{d\omega}{i}\alpha_i(\omega)\alpha_j(\omega)$$

De tal manera que esta integral ha servido para introducir el concepto de polarizabilidad dinámica. Por lo tanto, de las ecuaciones (52) y (53) se obtiene

$$W(R) \cong -\frac{3\hbar}{2R^6}\left(\frac{e^2 n_1}{m\omega_1^2}\right)\left(\frac{e^2 n_2}{m\omega_2^2}\right)\frac{\omega_1\omega_2}{\omega_1+\omega_2}$$

$$= -\frac{3\hbar}{2R^6}\alpha_1(0)\alpha_2(0)\frac{\omega_1\omega_2}{\omega_1+\omega_2} \qquad (54)$$

$$= -\frac{3\hbar}{2R^6}\frac{e}{m^{1/2}}\frac{\alpha_1(0)\alpha_2(0)}{\left[\alpha_1(0)/n_1\right]^{1/2}+\left[\alpha_2(0)/n_2\right]^{1/2}} \qquad (55)$$

La ecuación (54) es análoga a la ecuación (40). De las ecuaciones (54) y (55), la interacción puede estimarse en términos del valor estático de la polarizabilidad en aquellas

situaciones donde cada molécula está caracterizada por una fuerte frecuencia de absorción. La ecuación (55) la obtuvieron Slater y Kirkwood (1931).

La aproximación implicada en la ecuación (53) obviamente es equivalente a tratar cada molécula como oscilador armónico con una frecuencia ω_j y polarizabilidad estática $e^2 n_j \left/ m\omega_j^2 \right.$. Este modelo de oscilador de una molécula frecuentemente es muy conveniente para análisis cualitativo y semicualitativo de fuerzas entre moléculas. De hecho, London (1930) lo utilizó para derivar la ley ($1/R^6$) de fuerza de dispersión.

Otros aspectos de fuerzas intermoleculares vienen de la anisotropía del tensor de polarizabilidad de una molécula. Este efecto se puede examinar mejor con la interacción debida a momentos multipolares inducidos de orden superior (Blokhintsev, 1964; Margenau y Kestner, 1971). En el Anexo E se muestran brevemente la expansión multipolar.

Efecto de retardamiento

El efecto de retardamiento en la interacción electromagnética entre dos moléculas aparece cuando una de ellas emite una señal que no llega instantáneamente a la otra, más bien llega con un cierto retardo dependiendo de la separación entre ellas.

La fuerza de dispersión entre dos moléculas puede describirse como la inducción de polarización de una de ellas debido al campo de polarización instantáneo de la otra. La energía de interacción de la dispersión entre un par de moléculas a una distancia mayor que la longitud de onda de la radiación, por causa de transiciones dipolares en ellas, decae como ($1/R^7$) de acuerdo con la fórmula (Mahanty y Ninham, 1976):

$$E(R) = -\frac{23}{4\pi}\hbar c \frac{\alpha_1(0)\alpha_2(0)}{R^7}$$

En términos semiclásicos la energía de interacción puede definirse como el cambio en la energía del punto-cero de los modos de vibración del campo electromagnético cuando estas son perturbadas por las moléculas a través del acoplamiento del campo con las corrientes de polarización inducidas en las moléculas. La ecuación anterior es precisamente la que se debe derivar ahora para apreciar el efecto de retardamiento entre las señales emitidas por cada una del par de moléculas interactuantes.

Si una molécula está sujeta a un campo eléctrico variable

$$\mathbf{E}(t)=\mathbf{E}_0(\omega)\cos\omega t$$

el cual corresponde a la parte real de una onda plana. La polarización inducida en ella está dada por

$$\mathbf{P}(t)=\mathbf{P}(\omega)\cos\omega t=\left[\boldsymbol{\alpha}(\omega)\mathbf{E}_0(\omega)\right]\cos\omega t$$

en esta relación el tensor $\boldsymbol{\alpha}(\omega)$ tiene componentes

$$\alpha_{xy}(\omega)=-\frac{1}{\hbar}\sum_n\left(\frac{(\mathbf{d}_{\ell n})_y(\mathbf{d}_{n\ell})_x}{\omega_{n\ell}-\omega}+\frac{(\mathbf{d}_{\ell n})_y(\mathbf{d}_{n\ell})_x}{\omega_{n\ell}+\omega}\right),$$

donde el tensor $\mathbf{d}=-\sum e_j\mathbf{r}_j$ es el operador de momento dipolar de la molécula. La forma de $\boldsymbol{\alpha}(\omega)$ es derivada en muchos textos de mecánica cuántica.

Supóngase que las moléculas son esencialmente partículas puntuales, tal que en la presencia de un campo eléctrico $\mathbf{E}(\mathbf{r},t)$ la densidad de polarización inducida puede escribirse como

$$\mathbf{p}(\mathbf{r},t)=\left(\int_{-\infty}^{\infty}\boldsymbol{\alpha}(t-t')\mathbf{E}(\mathbf{R},t')dt'\right)\delta(\mathbf{r}-\mathbf{R}) \qquad (56)$$

donde $\mathbf{R}$ es la coordenada de posición de la molécula. Aquí $\alpha(t)$ es la transformada de Fourier de $\alpha(\omega)$ definida por

$$\boldsymbol{\alpha}(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\omega e^{i\omega t}\boldsymbol{\alpha}(\omega)$$

y $\mathbf{E}(\mathbf{r},t)$ puede escribirse en la forma

$$\mathbf{E}(\mathbf{r},t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\omega\,\mathbf{E}(\mathbf{r},\omega)e^{i\omega t}$$

La ecuación (56) se sigue del hecho de que la polarización manifestada sobre la molécula es la suma de aquella debida a cada componente de la frecuencia de Fourier del campo eléctrico. En general, $\alpha(\omega)$ puede definirse con una pequeña parte imaginaria en ω tal que su transformada de Fourier $\alpha(t)$ desaparece para $t < 0$, así como para satisfacer el

principio de causalidad. Entonces en la ecuación (56) el límite superior de la integración es t, y su transformada de Fourier puede escribirse como,

$$\mathbf{p}(\mathbf{r},\omega) = [\boldsymbol{\alpha}(\omega)\mathbf{E}(\mathbf{R},\omega)]\delta(\mathbf{r}-\mathbf{R})$$

donde $$\mathbf{p}(\mathbf{r},\omega) = \int_{-\infty}^{\infty} e^{-i\omega t}\mathbf{p}(\mathbf{r},t)dt \qquad (57)$$

Para considerar el efecto de acoplamiento entre un par de moléculas en $\mathbf{R}_1$ y $\mathbf{R}_2$ con polarizabilidad $\alpha_1(\omega)\ y\ \alpha_2(\omega)$, y el campo electromagnético, nótese que la corriente de polarización inducida sobre las moléculas es

$$\mathbf{j}(\mathbf{r},t) = \frac{\partial}{\partial t}\mathbf{p}(\mathbf{r},t)$$

y la correspondiente transformada de Fourier, definida como en la ecuación (57), satisface la relación

$$\begin{aligned}\mathbf{j}(\mathbf{r},\omega) &= i\omega\mathbf{p}(\mathbf{r},\omega)\\ &= i\omega[\boldsymbol{\alpha}_1(\omega)\mathbf{E}(\mathbf{R}_1,\omega)\delta(\mathbf{r}-\mathbf{R}_1)+\boldsymbol{\alpha}_2(\omega)\mathbf{E}(\mathbf{R}_2,\omega)\delta(\mathbf{r}-\mathbf{R}_2)]\end{aligned} \qquad (58)$$

Las ecuaciones de Maxwell del campo electromagnético ahora pueden escribirse con los términos de la fuente correspondientes a la corriente de desplazamiento $\mathbf{j}(\mathbf{r},t)$ (Jackson, 1962):

Ley de Coulomb $$\nabla\cdot\mathbf{D} = 4\pi\rho \qquad (59)$$

Ley de Ampere $$\nabla\times\mathbf{H} = \frac{4\pi}{\varepsilon}\mathbf{j} + \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}$$

Ley de Faraday $$\nabla\times\mathbf{E} + \frac{1}{c}\frac{\partial B}{\partial t} = 0$$

Ausencia de polos magnéticos libres $\nabla\cdot\mathbf{B} = 0$

Por esta última ecuación se puede definir **B** en términos de un potencial vectorial

$$\mathbf{B} = \nabla\times\mathbf{A} \qquad (60)$$

entonces la ley de Faraday puede escribirse como

$$\nabla\times\left(\mathbf{E} + \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}\right) = 0$$

esto significa que la cantidad entre paréntesis puede escribirse como el gradiente de alguna función escalar (Anexo F), por decir, un potencial escalar Φ; esto es

$$\mathbf{E}+\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}=-\nabla\Phi$$

o

$$\mathbf{E}=-\nabla\Phi-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \tag{61}$$

La definición de **B** y **E** en términos de los potenciales **A** y Φ satisface idénticamente las dos ecuaciones de Maxwell. La conducta dinámica de **A** y Φ será determinada por las dos ecuaciones inhomogéneas.

Aunque la especificación de **A** determina completamente a **B** de acuerdo con la ecuación (60), el caso contrario no es verdadero, puesto que el rotacional del gradiente de cualquier escalar desaparece idénticamente, así que podemos sumar a **A** el gradiente de un escalar ξ sin afectar **B**. Esto es, **A** puede ser reemplazado por

$$\mathbf{A}' = \mathbf{A} + \nabla\xi \tag{62}$$

Entonces (Marion, 1965)

$$\mathbf{E}=-\nabla\Phi-\frac{1}{c}\frac{\partial}{\partial t}\left(\mathbf{A}+\nabla\xi\right)$$

o

$$\mathbf{E}=-\nabla\left(\Phi+\frac{1}{c}\frac{\partial\xi}{\partial t}\right)-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}$$

por lo tanto, si se hace la transformación (62), también se debe reemplazar Φ por

$$\Phi' = \Phi - \frac{1}{c}\frac{\partial\xi}{\partial t} \tag{63}$$

para que la expresión de **E** quede inalterada. Las transformaciones (62) y (63) se llaman *transformaciones de norma*. Aun cuando sumemos el gradiente de una función escalar al potencial vectorial y sumemos la derivada del tiempo de esta función al potencial escalar, los vectores de campo quedan inalterados. Esto es, son las *cantidades de campo* y no los *potenciales* las que poseen significado físico. Por lo tanto, se dice que los campos vectoriales son invariantes a las transformaciones de norma, es decir, son de *norma invariante*.

Debido a la arbitrariedad en la elección de la norma, se está en libertad de imponer una condición adicional en **A**. Se puede argumentar esto, en otros términos: un vector no se especifica completamente solo con su rotacional y su divergencia para que quede únicamente determinado. Evidentemente se tiene la ventaja de hacer una elección de la $\nabla \cdot \mathbf{A}$ que simplifique el problema particular bajo consideración. Una posibilidad de uso frecuente es

$$\nabla \cdot \mathbf{A} = -\frac{1}{c}\frac{\partial \Phi}{\partial t} \tag{64}$$

Esta condición es la *condición de Lorentz* y especifica la *norma de Lorentz* de los potenciales. Aplicando las condiciones de norma (62) y (63) a la norma de Lorentz (64) se obtiene

$$\nabla \cdot \mathbf{A}' + \frac{1}{c}\frac{\partial \Phi'}{\partial t} = \nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t} + \left(\nabla^2 \xi - \frac{1}{c^2}\frac{\partial^2 \xi}{\partial t^2} \right)$$

si los potenciales $\mathbf{A}, \Phi \;\; y \;\; \mathbf{A}', \Phi'$ obedecen la condición de Lorentz, entonces se debe requerir que ξ satisfaga la ecuación de onda

$$\nabla^2 \xi - \frac{1}{c^2}\frac{\partial^2 \xi}{\partial t^2} = 0$$

Por simplicidad, este se restringe al caso especial del vacío, tal que $\varepsilon, \mu = 1$, puesto que los medios materiales solo agregan complejidad en sus fronteras e inhomogeneidades que no es deseable examinar en este momento. Entonces, la ley de Coulomb en el vacío se puede escribir como

$$\nabla \cdot \mathbf{E} = 4\pi\rho$$

sustituyendo la ecuación (61) en la ecuación anterior se obtiene

$$\nabla \cdot \left(-\nabla \Phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \right) = 4\pi\rho$$

$$\Rightarrow \quad \nabla^2 \Phi + \frac{1}{c}\frac{\partial}{\partial t}\nabla \cdot \mathbf{A} = -4\pi\rho \tag{65}$$

La ley de Faraday (de las ecuaciones de Maxwell) en el vacío, se puede escribir como

$$\nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j}$$

y sustituyendo las ecuaciones (60) y (61) en esta última ecuación se obtiene

$$\nabla\times\nabla\times\mathbf{A}+\frac{1}{c}\nabla\frac{\partial\Phi}{\partial t}+\frac{1}{c^2}\frac{\partial\mathbf{A}}{\partial t^2}=\frac{4\pi}{c}\mathbf{j}$$

usando la identidad vectorial $\nabla\times\nabla\times\mathbf{A}=\nabla(\nabla\cdot\mathbf{A})-\nabla^2\mathbf{A}$ se obtiene

$$\nabla^2\mathbf{A}-\frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2}-\nabla\left(\nabla\cdot\mathbf{A}+\frac{1}{c}\frac{\partial\Phi}{\partial t}\right)=-\frac{4\pi}{c}\mathbf{j} \tag{66}$$

Entonces las ecuaciones (65) y (66) dan una representación de las cuatro ecuaciones de Maxwell, usando solamente dos ecuaciones para los potenciales $A\ y\ \Phi$. Pero estas dos ecuaciones están acopladas (justo como estaban las ecuaciones de Maxwell); es decir, ambas ecuaciones involucran los dos potenciales. La aplicación de la condición de Lorentz (64) produce el desacoplamiento deseado:

$$\begin{aligned}\nabla^2\Phi-\frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2}&=-4\pi\rho\\ \nabla^2\mathbf{A}-\frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2}&=-\frac{4\pi}{c}\mathbf{j}\end{aligned} \tag{67}$$

Por lo tanto, se encuentra que en la norma de Lorentz los potenciales del campo satisfacen las ecuaciones de onda inhomogéneas.

Las transformadas de Fourier de las ecuaciones (67), (64) y (61) son

$$\left(\nabla^2+\frac{\omega^2}{c^2}\right)\mathbf{A}(\mathbf{r},\omega)=-\frac{4\pi}{c}\mathbf{j}(\mathbf{r},\omega) \tag{68}$$

$$\begin{aligned}\nabla\cdot\mathbf{A}(\mathbf{r},\omega)+\frac{i\omega}{c}\phi(\mathbf{r},\omega)&=0\\ \mathbf{E}(\mathbf{r},\omega)=-\frac{i\omega}{c}\mathbf{A}(\mathbf{r},\omega)-\nabla\phi(\mathbf{r},\omega)&\\ =-\frac{ic}{\omega}\left(\frac{\omega^2}{c^2}\mathbf{A}+\nabla(\nabla\cdot\mathbf{A})\right)&\end{aligned} \tag{69}$$

Al sustituir la ecuación (58) en (68) se obtiene

$$\left(\nabla^2+\frac{\omega^2}{c^2}\right)\mathbf{A}(\mathbf{r},\omega)=-\frac{4\pi i\omega}{c}\left[\boldsymbol{\alpha}_1\delta(\mathbf{r}-\mathbf{R}_1)+\boldsymbol{\alpha}_2(\omega)\delta(\mathbf{r}-\mathbf{R}_2)\right]\mathbf{E}(\mathbf{r},\omega) \tag{70}$$

substituyendo (69) en (70) se adquiere

$$\left(\nabla^2+\frac{\omega^2}{c^2}\right)\mathbf{A}(\mathbf{r},\omega)=-4\pi\left[\boldsymbol{\alpha}_1\delta(\mathbf{r}-\mathbf{R}_1)+\boldsymbol{\alpha}_2(\omega)\delta(\mathbf{r}-\mathbf{R}_2)\right]\left[\left(\frac{\omega^2}{c^2}\right)\mathbf{A}(\mathbf{r},\omega)+\nabla\{\nabla\cdot\mathbf{A}(\mathbf{r},\omega)\}\right]$$

puesto que el operador $\left[\left(\frac{\omega^2}{c^2}\right)\zeta+\nabla(\nabla\cdot)\right]$ conmuta con $\left[\nabla^2+\left(\frac{\omega^2}{c^2}\right)\right]$, entonces

$$\left(\nabla^2+\frac{\omega^2}{c^2}\right)\mathbf{E}(\mathbf{r},\omega)=-\frac{4\pi i\omega}{c}\left[\boldsymbol{\alpha}_1\delta(\mathbf{r}-\mathbf{R}_1)+\boldsymbol{\alpha}_2(\omega)\delta(\mathbf{r}-\mathbf{R}_2)\right]\mathbf{A}(\mathbf{r},\omega)$$

donde $\mathbf{A}(\mathbf{r},\omega)=-\frac{ic}{\omega}\left(\frac{\omega^2}{c^2}\mathbf{E}(\mathbf{r},\omega)+\nabla(\nabla\cdot\mathbf{E}(\mathbf{r},\omega))\right)$. Esto es

$$\left(\nabla^2+\frac{\omega^2}{c^2}\right)\mathbf{E}(\mathbf{r},\omega)=-4\pi\left[\boldsymbol{\alpha}_1(\omega)\delta(\mathbf{r}-\mathbf{R}_1)+\boldsymbol{\alpha}_2(\omega)\delta(\mathbf{r}-\mathbf{R}_2)\right]\left[\left(\frac{\omega^2}{c^2}\right)\mathbf{E}(\mathbf{r},\omega)+\nabla\{\nabla\cdot\mathbf{E}(\mathbf{r},\omega)\}\right]$$

$$=-4\pi\left[\boldsymbol{\alpha}_1(\omega)\left[\left(\frac{\omega^2}{c^2}\right)+\nabla_{\mathbf{r}}\nabla_{\mathbf{R}_1}\right]\mathbf{E}(\mathbf{R}_1,\omega)\delta(\mathbf{r}-\mathbf{R}_1)+\boldsymbol{\alpha}_2(\omega)\left[\left(\frac{\omega^2}{c^2}\right)+\nabla_{\mathbf{r}}\nabla_{\mathbf{R}_2}\right]\mathbf{E}(\mathbf{R}_2,\omega)\delta(\mathbf{r}-\mathbf{R}_2)\right]$$

Efectos del solvente

La interacción entre moléculas o partículas pequeñas en un medio solvente puede ser muy diferente de aquellas de moléculas aisladas en espacio libre o en un gas. La presencia de un medio suspensor simplemente reduce la energía de interacción o fuerza por un factor ϵ o ϵ^2. Primero, el momento dipolar intrínseco y la polarizabilidad de una molécula de gas aislado podría ser diferente en el estado líquido o cuando es disuelta en un medio. Esto depende de un camino complicado en sus interacciones con el solvente que lo rodea y usualmente solo puede ser encontrado por experimento. Segundo, una molécula disuelta solo puede moverse desplazando un volumen igual de solvente de su trayectoria; en consecuencia, la polarizabilidad α en un medio debe representar la *polarizabilidad de exceso* de una molécula o partícula sobre aquella del solvente y debe desaparecer cuando una partícula disuelta tiene las mismas propiedades que el solvente. Cualitativamente, se puede decir que, si ningún campo eléctrico es reflejado por una partícula, es 'invisible' en el medio solvente y consecuentemente no experimenta una fuerza.

a) Efecto en la constante dieléctrica

El problema de saber la polarizabilidad de exceso o efectiva puede ser planteado al tratar una molécula disuelta o una partícula pequeña como un medio dieléctrico separado de una forma y tamaño dados. Este modelo continuo tiene una ventaja obvia puesto que la constante dieléctrica de un medio es usualmente conocida. Acordadamente, una molécula *i* podría ser modelada como una esfera dieléctrica de radio a_i y constante dieléctrica ε_i. Ahora bien, en un medio de constante dieléctrica ε (figura 5.3a de Israelachvili (1985), tal esfera dieléctrica estará polarizada por un campo E y adquirirá un momento dipolar de exceso dado por (Landau y Lifshitz, 1963)

$$p_{ind} = 4\pi\varepsilon_0\varepsilon\left(\frac{\varepsilon_i - \varepsilon}{\varepsilon_i + 2\varepsilon}\right)a_i^3 E$$

así que su polarizabilidad efectiva o de *exceso* en el medio es

$$\alpha_i = 4\pi\varepsilon_0\varepsilon\left(\frac{\varepsilon_i - \varepsilon}{\varepsilon_i + 2\varepsilon}\right)a_i^3 = 3\varepsilon_0\varepsilon\left(\frac{\varepsilon_i - \varepsilon}{\varepsilon_i + 2\varepsilon}\right)\upsilon_i$$

donde $\upsilon_i = \frac{4}{3}\pi a_i^3$ es el volumen de la molécula o esfera. Esta ecuación muestra que para una esfera dieléctrica de alta ε_i en espacio libre (donde ε = 1) su polarizabilidad es aproximadamente $\alpha_i \approx 4\pi\varepsilon_0 a^3$, como se encontró previamente para un átomo simple de un electrón. Además, si $\varepsilon > \varepsilon_i$, la polarizabilidad es negativa, indicando que la dirección del dipolo inducido es opuesta a aquella en espacio libre.

b) Capas de hidratación o zonas de solvatación

Para iones pequeños o multivalentes en solventes altamente polares, la fuerte dependencia de la orientación de su interacción ion-dipolo tenderá a orientar las moléculas de solvente alrededor de ellos, favoreciendo $\theta = 0°$ cerca de los cationes y θ = 180° cerca de los aniones. Entonces, iones de Li^+, Be^{2+}, Mg^{2+} y Al^{3+} en agua tienen un número de moléculas de agua orientacionalmente ligadas a ellas. Tales iones son llamados *iones solvatados* o *iones hidratados*, y el número de moléculas de agua que se amarran —usualmente entre 4 y 6— es

conocido como su *número de hidratación*. Debería notarse, sin embargo, que estas moléculas de agua no están completamente inmovilizadas y que se intercambian con agua en el bulto, aunque más lentamente. El número de hidratación es un indicador cualitativo del grado al que los iones amarran agua más bien que un valor exacto.

Cercanamente relacionado al número de hidratación está el radio efectivo o *radio hidratado* de un ion en agua, que es más grande que su radio real (*i.e.* su radio en la red cristalina). Debido a que los iones más pequeños son más hidratados tienden a tener radio hidratado mayor que iones más grandes. Números y radio de hidratación pueden ser deducidos de mediciones de viscosidad, difusión, compresibilidad, conductividad, solubilidad, y varias propiedades termodinámicas y espectroscópicas de soluciones electrolíticas, los resultados raramente están de acuerdo entre sí.

Más intuición en la naturaleza de la hidratación del ion puede ganarse considerando el tiempo promedio que moléculas de agua permanecen ligadas a iones. En líquido puro a temperatura ambiente las moléculas de agua caen alrededor con un tiempo medio de reorientación o tiempo de correlación rotacional de cerca de 10^{-11} s. Esto también da una estimación del tiempo de vida de la ligadura agua-agua formada en agua líquida (las ligaduras de hidrógeno). Pero cuando las moléculas de agua están cerca de iones, varias técnicas tales como resonancia magnética nuclear de oxígeno, muestran que las vidas medias o tasa de intercambio de moléculas de agua en la primera capa de hidratación puede ser mucho más grande, variando de 10^{-11}s a muchas horas. Para iones solvatados muy débilmente (usualmente iones monovalentes grandes) tales como $N(CH3)_4^+$, Cl^-, Br^- y I^-, estas vidas medias no son muy diferentes de aquellas para dos moléculas de agua, y aun pueden ser más cortas (referidas como *hidratación negativa*).

Interacciones moleculares electrostáticas

Las fuerzas dipolo-dipolo son alrededor del 1 % tan fuertes como las ligaduras covalentes o iónicas, y se hacen débiles conforme la distancia entre los dipolos aumenta (Zumdahl y Decoste, 2014). En la fase gas estas fuerzas son relativamente importantes, dado que las moléculas están alejadas.

Particularmente las fuerzas potentes dipolo-dipolo ocurren entre moléculas en las que el hidrógeno está ligado a un átomo altamente electronegativo, tal como nitrógeno, oxígeno,

o flúor. Dos factores importantes para las intensidades de estas interacciones son la gran polaridad de la ligadura y la cercanía de los dipolos, la cual se hace posible por el tamaño tan pequeño del hidrógeno.

Algunos enlaces inusualmente fuertes y dependientes de la orientación están involucrados en interacciones entre moléculas de agua. Se sabe que la parte enlazante del enlace es responsable de esta interacción, y se observa midiendo las distancias entre varios centros atómicos en una red de hielo en el caso del agua. La distancia intramolecular O–H es de aproximadamente 0.10 nm para este enlace covalente, pero la distancia intermolecular O·H es de 0.176 nm, mucho menos de 0.26 nm correspondiente a la suma de los dos radios de Van der Waals, pero aún mayor que la distancia covalente de 0.10 nm. Dichos enlaces se conocen como enlaces de hidrógeno (Israelachvili, 1985). Los enlaces de hidrógeno no son exclusivos del agua; existen en diversos grados entre átomos electronegativos (p. ej., O, N, F y Cl) y átomos H de hidrógeno unidos covalentemente a átomos electronegativos similares. Estos enlaces solo involucran átomos de hidrógeno, que pueden interactuar fuertemente con los átomos electronegativos cercanos, lo que resulta en un *enlace* mediado por H efectivo entre dos átomos electronegativos. El enlace de hidrógeno es predominantemente una interacción electrostática (coulómbica). El átomo H no se comparte, sino que permanece más cerca y unido covalentemente a su átomo padre. En consecuencia, el enlace de hidrógeno entre dos grupos XH e Y se denota por X-H···Y. Sin embargo, ciertas características de los enlaces de hidrógeno hacen que aparezcan como enlaces covalentes débiles, no solo son bastante fuertes sino también bastante direccionales, de modo que forman *estructuras* tridimensionales débiles en los sólidos, mientras que en los líquidos el orden de corto alcance puede dar lugar a líquidos asociados. La fuerza de la mayoría de los enlaces de hidrógeno se encuentra entre 10 y 40 kJ mol^{-1} por enlace a temperatura ambiente, lo que los hace más fuertes que un enlace típico de Van der Waals (~ 1 kJ mol^{-1}) pero aún más débil que los enlaces covalentes o iónicos (~ 500 kJ mol^{-1}). Aunque el enlace de hidrógeno sea una interacción electrostática, encontrar el potencial de interacción, es una tarea ardua en cada caso. Las fuerzas de enlaces de hidrógeno tienen una dependencia de la distancia que va como $1/r^2$, y es la misma que para la interacción ion-dipolo, dada por la ecuación (31), y ahora expresada como

$$W(r,\theta) = {-Q_{H^+}\, p\cos\theta}\Big/{4\pi\varepsilon\varepsilon_0 r^2}$$

La magnitud de esta carga Q_{H^+} no se conoce de antemano. Por ejemplo, el enlace-H de agua-agua con Q = 0.24e = 0.4 × 10^{-19} C, p_{H2O}=1.5 D = 5 × 10^{-30} Cm (el momento dipolar del agua), r = 0.176 + (0.100/2) = 0.226 nm (la distancia entre H^+ y el centro del dipolo O^-H^+), θ=0, and ε = 1, obtenemos $W(r,\theta) = 3.4 \times 10^{-20}$J = 20.4 kJ mol^{-1} = 4.9 kcal/mole. Esto concuerda con el valor de literatura de 4.5 kcal/mol para el enlace H agua-agua. Los enlaces de hidrógeno pueden ocurrir intermolecularmente e intramolecularmente y pueden existir en ambiente no polar. Estos son importantes tanto en ensamblajes macromoleculares y biológicos, como en proteínas que unen diferentes segmentos dentro de las moléculas y en ácidos nucleicos responsables de la estructura de la molécula de ADN. Su participación en la creación de estructuras macromoleculares monodimensionales, bidimensionales y tridimensionales se denomina polimerización con enlaces de hidrógeno.

Puentes de hidrógeno

Una situación inusual en la que átomos no-ligados quedan cercanamente juntos, está en los puentes de hidrógeno, los cuales son un tipo especial de atracción dipolo-dipolo. Se forman más frecuentemente cuando un átomo de hidrógeno que está covalentemente ligado a un átomo de nitrógeno, de oxígeno, o de flúor interactúa con el único par de electrones de un segundo de tales átomos cercanos. Los puentes de hidrógeno son más fuertes que otras atracciones dipolo-dipolo, pero aún son mucho más débiles que las ligaduras covalentes.

Por ejemplo, moléculas de agua en la fase gas forman un dímero $(H_2O)_2$ débilmente asociado por medio de puentes de hidrógeno. Un átomo de una molécula de agua penetra levemente la *orilla dura* del átomo de oxígeno de otra. Este átomo de hidrógeno mantiene una ligadura covalente regular con un átomo de oxígeno, pero forma una ligadura más débil (puente de hidrógeno) y más larga con el segundo átomo de oxígeno. Descansa casi exactamente sobre una línea recta que une los dos átomos de oxígeno y queda significativamente más cerca al segundo oxígeno de lo que se esperaría de las fuerzas repulsivas. En el puente de hidrógeno H_2O-H_2O, el átomo de hidrógeno tiene una carga positiva fraccional, y es atraído a la carga negativa fraccional en la vecindad del átomo de oxígeno, como se esperaría en cualquier atracción dipolo-dipolo. Sin embargo, la interacción es más fuerte que otras atracciones dipolo-dipolo por dos razones. i) La primera razón es que el hidrógeno en el puente polar O–H está protegido levemente por electrones debido a que el

hidrógeno no tiene otros electrones que los de valencia. Como resultado, puede aproximarse a pares de electrones solos sobre el átomo de oxígeno de la segunda molécula, llevando a grandes interacciones electrostáticas entre las dos. ii) La segunda razón es que una pequeña cantidad de ligaduras covalentes viene del comportamiento de electrones entre los dos átomos de oxígeno y el átomo de hidrógeno que interviene. La ligadura covalente explica porque el átomo de hidrógeno penetra el átomo de oxígeno. Estos detalles son típicos de muchos puentes de hidrógeno.

Los puentes de hidrógeno pueden ser intermoleculares (entre moléculas) o intramoleculares (entre átomos de la molécula) y forman las estructuras de muchas substancias biológicas importantes y complejas (Oxtoby et al., 1990); los primeros son importantes en las estructuras del estado líquido y estado sólido de muchas substancias simples.

Interacción hidrofóbica

La interacción hidrofóbica describe una fuerte atracción entre moléculas hidrofóbicas y superficies en agua, a menudo más fuerte que su atracción en el espacio libre. Por ejemplo, la energía de interacción de Van der Waals entre dos moléculas de metano en contacto en el espacio libre es -2.5×10^{-21} J mientras que en agua es -14×10^{-21} J. Análogamente, las tensiones superficiales de la mayoría de los hidrocarburos saturados se encuentran en el rango de 15–30 mJ m^{-2}, mientras que para el agua están en el rango de 40–50 mJ m^{-2}. Debe quedar claro que no hay un enlace asociado con este fenómeno entrópico, que surge de la reorganización de las configuraciones de enlaces H en las zonas superpuestas de solvatación, ya que dos especies hidrofóbicas se juntan y forman un intervalo mucho mayor que cualquier enlace típico.

Ha habido pocas mediciones de interacción hidrofóbica entre moléculas no polares disueltas en agua, esencialmente porque son muy insolubles. El problema teórico de la interacción hidrofóbica entre dos moléculas es más complejo involucra otras moléculas y es de un intervalo más grande que el que surge de cualquier potencial simple aditivo por pares. Las conectividades de enlace-H de la red y el papel del salto de protones en la modificación de la polarizabilidad local efectiva del agua deben incluirse en cualquier simulación por computadora. Las interacciones hidrofóbicas entre superficies extendidas se miden por

medio de la ley de fuerza hidrofóbica (Israelachvili y Pashley, 1982) entre dos superficies hidrofóbicas curvas macroscópicas en el agua en el rango de 4 a 10 nm, en el cual la fuerza decae exponencialmente con la distancia. Para moléculas de soluto pequeñas, el potencial por pares hidrófobo $W_H(r)$ es exponencial, proporcional al diámetro de las moléculas o grupos moleculares σ, y puede expresarse como

$$W_H(r) \approx -20\sigma e^{-(r-\sigma)/D_H} \qquad kJ\ mol^{-1}$$

donde σ está en nm y donde D_H es la longitud de descomposición hidrofóbica característica ~1.0 nm. La energía libre de dimerización implica poner $r = \sigma$ en las ecuaciones anteriores

$$\Delta G_{transfer} = W_H(r) \approx -20\sigma \text{ kJ mol}^{-1}$$

La fuerza hidrofóbica para separar dos moléculas del contacto está dada por

$$F_H(\sigma) = -(dW_H/dr)_{r=\sigma} = -3 \times 10^{-11}\sigma(nm) \qquad N\ a\ temp.amb.$$

Para $\sigma \approx 0.5$ nm se esperaría una fuerza de adhesión de ~15 pN. La interacción hidrófoba tiene un rango de 1.5–2.0 nm para moléculas como metano, ciclohexano y benceno. Esto es del orden del radio de las moléculas o grupos moleculares. Este largo rango tiene implicaciones para comprender la dinámica de las moléculas anfifílicas autoensambladas en micelas y bicapas, el plegamiento de proteínas, la agregación hidrofóbica y mecanismos de fusión de bicapas y membranas biológicas.

Otros aspectos de interacciones hidrofóbicas entre superficies extendidas y partículas coloidales se refieren a si se trata de una fuerza estructural (debido a la estructura del agua o una fuerza *de agotamiento*), de una atracción mejorada de Van der Waals (dados los efectos de salto de protones o de correlación de carga) o de una fuerza capilar debido al puenteo de burbujas de vapor.

Interacciones hidrofílicas

Los efectos hidrofílicos (o interacción hidrofílica) se reconocen tanto por la propensión de ciertas moléculas y grupos a ser solubles en agua como por repelerse fuertemente en el agua, en contraste con la fuerte atracción exhibida por grupos hidrofóbicos. Los grupos hidrofílicos (es decir, amantes del agua) prefieren estar en contacto con el agua en lugar de estar juntos, y a menudo son higroscópicos (absorben el vapor del agua). Algunas redes de polímeros hidrofílicos pueden hincharse en agua hasta 1000 veces su tamaño original, formando hidrogeles. Los iones fuertemente hidratados y los iones híbridos son hidrófilos. Algunas

moléculas no cargadas e, incluso, no polares pueden ser hidrófilas si tienen la geometría correcta, y si contienen átomos capaces de unirse en H con agua, por ejemplo, los átomos de oxígeno en grupos de alcohol y de óxido de etileno, y los átomos de nitrógeno en aminas.

Las diferencias entre superficies hidrofílicas y las superficies hidrofóbicas son sutiles pero importantes para comprender por qué las interacciones entre estas son tan diferentes. Una atracción *normal* de Van der Waals con una repulsión estérica es de corto alcance entre dos superficies hidrofílicas, mientras que una atracción hidrofóbica mejorada se obtiene entre dos superficies hidrofóbicas. El fuerte enlace direccional del agua a una superficie hidrofílica o molécula de soluto agrega una barrera estérica adicional, aumentando el rango de la fuerza repulsiva o el tamaño efectivo de la molécula.

La orientación de todas las moléculas de agua en una determinada dirección restringe parcialmente los procesos de transporte de carga. No se produce dicho enlace direccional a una interfaz hidrofóbica o de vapor, donde las moléculas de agua estén más libres de orientarse como quieran. Esto permite el transporte de carga y el salto de protones en el líquido en el bulto y, a su vez, da pie a la interacción hidrófoba impulsada por entropía. Se pueden identificar tres tipos de orden: i) ordenamiento posicional en capas, que afecta las fuerzas oscilatorias; ii) ordenamiento orientativo, que afecta las interacciones de transferencia de carga electrostática y entrópica; y iii) variaciones de densidad media cerca de las superficies, que pueden dar lugar a la repulsión estérica adicional entre las superficies hidrofílicas y una atracción de agotamiento entre las superficies hidrofóbicas.

Cuando se disuelven en agua, ciertas moléculas tienen un efecto drástico en otras moléculas de soluto. Se cree que esto se debe a su efectividad para alterar la estructura del agua local. Por ejemplo, cuando la urea, $(NH_2)_2C{=}O$, se disuelve en agua puede hacer que las proteínas se desplieguen y clasifiquen dichos compuestos no iónicos, pero altamente potentes en formar o romper estructuras. Dichas moléculas se conocen comúnmente como agentes caotrópicos, lo cual significa que su interrupción de la estructura del agua local conduce al caos. Parece entonces, que las interacciones hidrofílicas e hidrofóbicas, a diferencia de las coulómbicas y las de dispersión, son interdependientes y no aditivas. La hidrofilia de algunos grupos hidrofílicos, por ejemplo el grupo $-OH$, se puede neutralizar por completo cuando se unen a una cadena alquílica larga como $-(CH_2)_{11}CH_3$. Por el contrario,

la energía hidrofóbica por grupo CH_2 de una cadena de alcanos se reduce mucho cuando un grupo de cabeza hidrofílica se une al final de la cadena.

Potenciales de interacción

Como la fuerza es igual a menos el gradiente del potencial de la distancia relativa entre las distancias *i* y *j*, una descripción adecuada del sistema en términos de potenciales clásicos queda indicada con la siguiente ecuación:

$$\boldsymbol{F}_{ij} = -\frac{dV(r_{ij})}{dr_{ij}}\frac{\boldsymbol{r}_{ij}}{r_{ij}}$$

De la fidelidad con que se represente la interacción entre las partículas en una dinámica molecular depende la calidad de los resultados. La obtención de potenciales de interacción se logra ajustando datos experimentales, o bien llevando a cabo cálculos cuánticos *ab initio*.

Considerando el caso de un sistema que contiene *N* átomos, la energía potencial se puede dividir en términos que dependen de las coordenadas de átomos individuales, por pares, por tripletas, etc.

$$V = \sum_i v_1(\boldsymbol{r}_i) + \sum_i \sum_{j>i} v_2(\boldsymbol{r}_i, \boldsymbol{r}_j) + \sum_i \sum_{j>i} \sum_{k>j>i} v_3(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) + \cdots$$

El primer término: v_1 representa el efecto de un campo externo del sistema mientras que el segundo término: v_2, el potencial por pares depende de la magnitud del par $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$, por lo que se puede escribir $v_2(\boldsymbol{r}_{ij})$. Este puede realizarse con cálculos paso a paso (*single point*) para construir el pozo de potencial correspondiente. El tercer término: v_3 involucra tercias y puede resultar significativo en densidades de líquidos. Los términos de tres cuerpos en el potencial raramente se incluyen en simulaciones por computadora. La aproximación por pares da una buena descripción de propiedades de líquidos, ya que los efectos promedio de tres cuerpos se pueden incluir definiendo un potencial par *efectivo*, es decir:

$$V \approx \sum_i v_1(\boldsymbol{r}_i) + \sum_i \sum_{j>i} v_2^{eff}(\boldsymbol{r}_{ij})$$

Los potenciales pares que aparecen en simulaciones de computadora generalmente se tratan como potenciales pares efectivos de este tipo, representando todos los efectos de muchos-cuerpos; por simplicidad se usa la notación $v(\boldsymbol{r}_{ij})\ o\ v(r)$. Este potencial debe reproducir datos experimentales para poder confiar en él, y que se pueda volver a depender de la densidad, la temperatura, etcétera, mientras que el verdadero potencial de dos cuerpos $v_2(\boldsymbol{r}_{ij})$, no.

El potencial más simple y usado en simulaciones moleculares es el de Lennard-Jones 12-6, el cual es el caso particular más simple de la función de energía del potencial intermolecular de Mie(*n*,*m*), que tiene dos parámetros para ajustar:

$$v^{LJ}(r) = 4\varepsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right)$$

El potencial de Lennard-Jones (LJ) proporciona una descripción razonable de las propiedades de argón, vía simulación por computadora, si los parámetros $\varepsilon\ y\ \sigma$ son elegidos apropiadamente. El potencial tiene una cola atractiva de largo alcance de la forma $-1/r^6$, un pozo negativo de profundidad ε y una pared repulsiva con pendiente casi vertical a distancias menores que $r{\sim}\sigma$. La profundidad del pozo frecuentemente se pone como cociente en unidades de temperatura ε/k_B, donde k_B es la constante de Boltzmann, valores de $\varepsilon/k_B \approx 120K$ y $\sigma \approx 0.34nm$ proporcionan acuerdos razonables con propiedades experimentales de argón líquido.

Otros potenciales usados para simular líquidos son el potencial de esferas duras, el potencial de esferas suaves, el RLJ, etc.

Un modelo de potencial

$$U_p(r) = U_{intra}(r,\theta,\phi) + U_{inter}(r) \tag{71}$$

Donde U_{intra} es la energía intramolecular que representa la interacción entre átomos de una misma molécula (especie), mientras que U_{inter} es la energía intermolecular de interacción entre moléculas distintas. El primer término de la ecuación (71) se expresa como:

$$U_{intra}(r,\theta,\phi) = U_D(r_{ij}) + U_A(\theta_{ijk}) + U_{AD}(\phi_{ijkl}) \tag{72}$$

donde $U_D(r_{ij})$ y $U_A(\theta_{ijk})$ representan la energía potencial debida a las distancias y ángulos de enlace, respectivamente. El término restante, $U_{AD}(\phi_{ijkl})$, expresa la energía potencial debida a la rotación del ángulo de torsión.

El segundo término del lado derecho de la ecuación (71) es la energía intermolecular que está representada por los potenciales de LJ y de Coulomb:

$$U_{inter}(r) = U_{LJ}(r_{ij}) + U_{Coulomb}(r_{ij}) \tag{73}$$

El potencial intermolecular proporciona las contribuciones dentro de una molécula. Los grados de libertad son distancias de enlace, ángulos de enlace y ángulos de torsión.

Tanto las distancias de enlace como los ángulos de enlace se pueden modelar con funciones de tipo armónico

$$U(r_{ij}) = \frac{k_r}{2}(r_{ij} - r_0)^2 \tag{74}$$

$$U(\theta_{ijk}) = \frac{k_\theta}{2}(\theta_{ijk} - \theta_0)^2 \tag{75}$$

donde r_{ij} es la distancia entre los átomos *i* y *j*, r_0 es la distancia de equilibrio y k_r es la constante de enlace; θ_{ijk} es el ángulo entre los átomos *i*, *j* y *k*, θ_0 es el ángulo de equilibrio y k_θ es la constante del enlace.

La forma más simple de potencial $U(\theta_{ijk})$ de los ángulos de torsión es una función periódica en el ángulo de torsión asociado a los cuatro átomos involucrados (*i*, *j*, *k*, *l*). el potencial que describe este tipo de interacción en un modelo de energía potencial para etanolaminas está dado por:

$$U(\phi_{ijk}) = \sum_{n=0}^{5} C_n cos^n(\phi_{ijk}) \tag{76}$$

donde ϕ_{ijk} es el ángulo de torsión entre los átomos *ijkℓ*como se muestra en la figura 26.

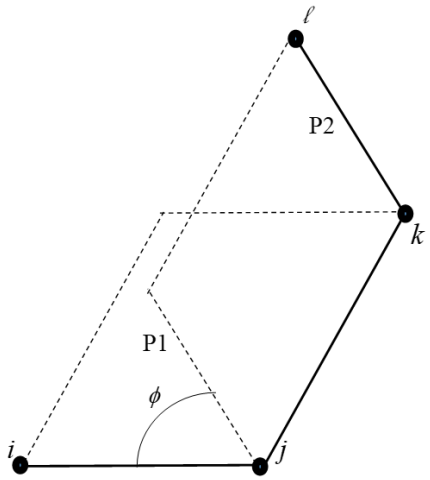
Para obtener las constantes C_i de las torsiones usadas en la simulación se ajustan las superficies de energía potencial obtenidas mediante cálculos *ab initio* en la ecuación (76).

Los ángulos que se refieren a la planaridad de un sistema de cuatro átomos en los que el central está enlazado a los otros tres, se conocen como ángulos de torsión impropios. Estos se representan por una función de tipo armónico:

$$U(\phi_{ijk\ell}) = \frac{k_\phi}{2}(\phi_{ijk} - \phi_0)^2 \tag{77}$$

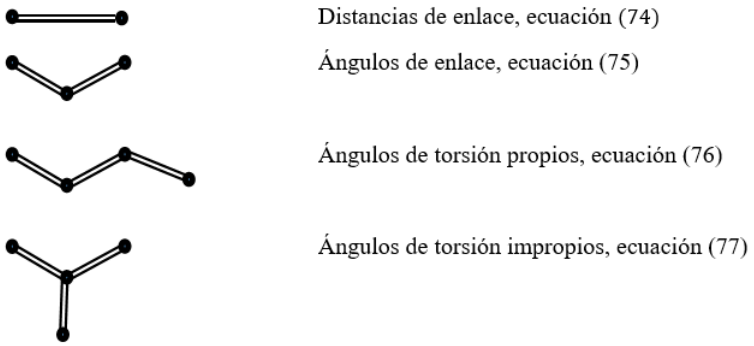
donde k_ϕ es la constante de enlace y ϕ_0 el ángulo de torsión de equilibrio. En la figura 27 se bosqueja cómo se consideran la distancia de enlace, los ángulos de enlace, ángulos de torsión propios y ángulos de torsión impropios.

Figura 26. *Ángulo diedro ϕ de la intersección de dos planos conformados por tres átomos, la torsión es referida a los átomos j y k*



Fuente: elaboración propia.

Figura 27. *Distancia y ángulos de acuerdo con las ecuaciones (74-77)*



Fuente: elaboración propia.

Potenciales intermoleculares

Son aquellos en los cuales cada átomo de una molécula interacciona con los de otra molécula con un potencial con dos contribuciones principales: corto alcance y largo alcance. Los potenciales de Lennard-Jones y de Coulomb las toman en cuenta.

Potenciales de corto alcance	**Potenciales de largo alcance**
Potencial de Mié: $$V(r) = \frac{C_n}{r^n} - \frac{C_m}{r^m}$$ Potencial de Lennard-Jones: $$U(r_{iajb}) = 4\epsilon\left[\left(\frac{\sigma}{r_{iajb}}\right)^{12} - \left(\frac{\sigma}{r_{iajb}}\right)^{6}\right]$$ O bien $$V(r) = \frac{A}{r^{12}} - \frac{B}{r^6}$$ donde $A = 4\epsilon\sigma^{12}$ $\quad B = 4\epsilon\sigma^6$ $$\sigma = \left(\frac{A}{B}\right)^{1/6} \qquad \epsilon = \frac{B^2}{4A}$$ Un truncamiento usual es $r_c = 2.5$	Potencial de Coulomb: $$U_C = \frac{1}{4\pi\epsilon_0}\frac{q_{ia}q_{jb}}{r_{iajb}}$$ Sumas de Ewald: $$U = \frac{1}{4\pi\epsilon_0}\frac{1}{2}\sum_n{}'\left(\sum_{i=1}^{n}\sum_{j=1}^{n} q_i\, q_j\, \left|\boldsymbol{r}_{ij} + n\right|^{-1}\right)$$ con potencial de Coulomb $$V(r) = f\frac{erfc(\beta r_{ij})}{r_{ij}}q_i q_j$$

En sistemas multicomponentes la interacción es entre el átomo *a* en la molécula *i* y el átomo *b* en la molécula *j*, σ es una medida del diámetro de los átomos que interaccionan y ϵ es un parámetro que representa la medida de la atracción entre los sitios de moléculas distintas, y r_{iajb} es la distancia entre el sitio *a* en la molécula *i* y el sitio *b* en la molécula *j*. Los parámetros σ y ϵ se ajustan a datos experimentales. Adicionalmente, ϵ_0 es la permitividad en el vacío y r_{iajb} es la distancia relativa entre la carga q_{ia} del átomo *a* en la molécula *i* y la carga q_{jb} del átomo *b* en la molécula *j*.

El potencial de Lennard-Jones describe interacciones entre átomos de líquidos monoatómicos y las interacciones de Van der Waals entre átomos de moléculas distintas. Las interacciones coulombianas decaen lentamente con la distancia entre las partículas, y alcanza valores aun en posiciones muy alejadas entre ellas, con lo que el potencial es mayor que la mitad de la arista de las cajas habituales de simulación y no es posible hacer un truncamiento del potencial. Un método eficiente para el cálculo de interacciones electrostáticas en un sistema periódico infinito es conocido como sumas de Ewald.

La esencia del potencial de interacción está en su capacidad para inferir el tipo de interacciones entre partículas, es decir, qué tipos de enlaces se tienen: covalente, iónico, metálico o Van der Waals. Es enlace covalente cuando los átomos comparten pares de electrones y se producen fuerzas atractivas. Es enlace iónico cuando un átomo con energía

de ionización baja pierde electrones fácilmente e interactúa con otro átomo que gana electrones fácilmente. En un enlace metálico los electrones de valencia de los átomos que comprenden un metal son comunes al agregado metálico completo tal que queda un tipo de *gas* de electrones libres. Las interacciones entre este gas de electrones y los iones metálicos positivos llevan a una fuerza cohesiva fuerte. El enlace tipo Van de Waals se exhibe en interacciones de corto alcance débiles entre átomos y moléculas debidas a fuerzas de Van der Waals o de dispersión.

Con esta información resulta muy débil el uso de un potencial fijo para describir el comportamiento de átomos y moléculas. Sin embargo, la utilidad de estos potenciales ha quedado ampliamente manifestada en su uso en modelos de mecánica clásica, como aquellos obtenidos en cálculos de dinámica molecular.

Cálculos mecánico-cuánticos de fuerzas intermoleculares

La fuerza siempre se puede escribir como el gradiente de un escalar más el rotacional de un vector

$$\boldsymbol{F} = -\nabla V + \nabla \times \boldsymbol{A}.$$

En un campo conservativo desaparece el rotacional del campo vectorial si y solo si el campo es menos el gradiente del potencial; en nuestro caso el campo vectorial es la fuerza $\boldsymbol{F}$, es decir

$$\nabla \times \boldsymbol{F} = 0 \Leftrightarrow \boldsymbol{F} = -\nabla V$$

Si la divergencia del campo vectorial desaparece, entonces $\boldsymbol{F}$ se puede expresar como el rotacional de un vector potencial ($\boldsymbol{A}$):

$$\nabla \cdot \boldsymbol{F} = 0 \Leftrightarrow \boldsymbol{F} = \nabla \times \boldsymbol{A}.$$

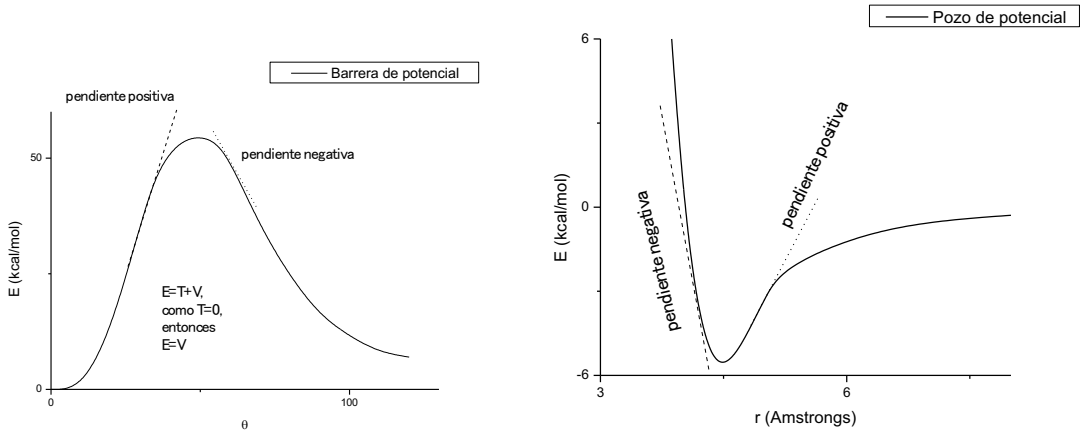
A partir de que la fuerza es el negativo del gradiente de la energía potencial, es decir

$$\boldsymbol{F} = -\nabla V$$

Las fuerzas entre moléculas dependen de la orientación en que cada átomo de una molécula interactúa con cada átomo de la otra molécula. La energía potencial más fuerte (repulsiva o atractiva) resulta cuando un átomo de una molécula se enfrenta directamente a un átomo de la otra molécula. Si la interacción es sesgada la fuerza de interacción resulta ser menor.

Conforme la interacción entre los átomos es más cercana la fuerza va pasando de atractiva a repulsiva. Esto se ilustra con barreras y pozos de potencial en la figura 28.

Figura 28. *Curvas de potencial: a) barrera de potencial, b) pozo de potencial*



a *b*

Fuente: elaboración propia.

Una pendiente positiva $\Longrightarrow \boldsymbol{\frac{dV}{d\theta}} > 0$ y $F = -\frac{dV}{d\theta}$ $\therefore F < 0$ indica una fuerza atractiva.

Una pendiente negativa $\Longrightarrow \boldsymbol{\frac{dV}{d\theta}} < 0$ y $F = \frac{dV}{d\theta}$ $\therefore F > 0$ indica una fuerza repulsiva.

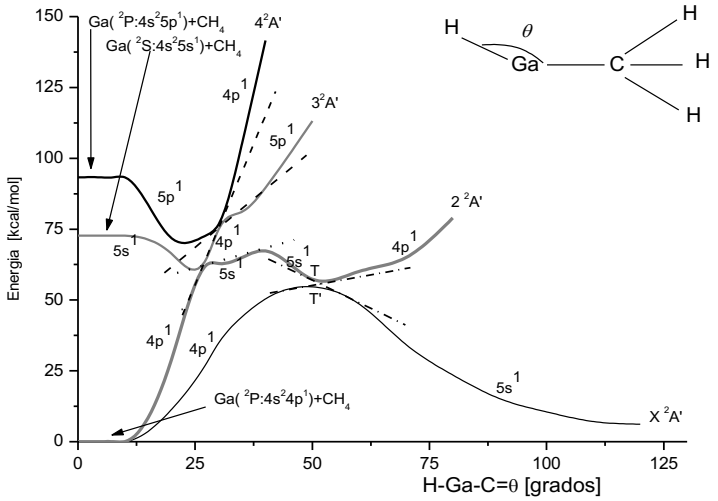
La energía total es la suma de las energías cinética + potencial; sin embargo, en cálculos *single point* la energía cinética es cero, por tanto, la energía total es exactamente igual a la energía potencial. Es de este modo que se construyen curvas de energía potencial o superficies de energía potencial, atacando a un objetivo fijo con un objetivo que se traslada punto a punto. Los objetivos pueden ser átomos o moléculas.

Interacción átomo-molécula

Un ejemplo del ataque de un átomo de metal a una molécula de gas es el caso de la interacción entre galio y metano. Son tres casos los que se han estudiado en esta interacción. Uno es el ataque de galio en su estado base Ga(^{2}P:4s^{2}4p^1) hacia un enlace C-H de metano (CH_4), con lo cual se genera una barrera de potencial. Otro es el ataque de galio en su primer estado excitado Ga(^{2}S:4s^{2}5s^1) hacia el mismo enlace C-H del metano (CH_4). Y el otro es el ataque

de galio en su segundo estado excitado Ga(^{2}P:4s^{2}5p^1) hacia el enlace C-H del metano (CH_4). Bajo la consideración de que estados de la misma especie no se cruzan, los resultados se muestran en la figura 29.

Figura 29. *Trayectorias de potencial del estado base y de los tres estados excitados mas bajos*



Fuente: elaboración propia.

La curva de energía potencial del doblete ^{2}P del Galio en su segundo estado excitado interactúa con metano en $\theta = 0$ con una energía de 93.3 kcal mol^{-1}, mientras que aquella del doblete ^{2}S del Galio en su primer estado excitado interactúa con metano en $\theta = 0$ con una energía de 72.7 kcalmol^{-1}. Nótese que en esta interacción resultó necesario recurrir a dos curvas del estado base para poder vencer las barreras de potencial y así encontrar la inserción del átomo de galio en un enlace C-H del metano, la cual se logra con cruces evitados de la trayectoria de reacción 3^{2}A' a la 2^{2}A', y de la 2^{2}A' a la X^{2}A' a través de las rectas de cruzamiento (líneas punteadas en la figura 29) en los puntos críticos, dada la naturaleza no-adiabática de las probabilidades de transición (Pacheco-Sánchez, Castillo, et al., 2007; Pacheco-Sánchez, Luna-García, et al., 2004).

El material correspondiente a las interacciones de un átomo de metal con una molécula se resumió en un artículo de revisión de artículos (Novaro et al., 2012). En este se llevó a cabo una revisión de estudios *ab initio* fundamentados en mecánica cuántica acerca

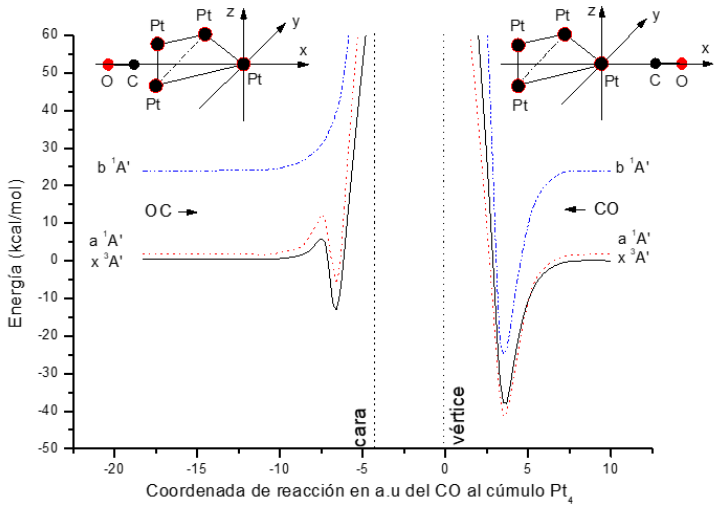
de mecanismos de reacción que conducen a la inserción de un átomo de metal en enlaces C – H, Si – H y Ge – H de metano, silano y germano, respectivamente, que finalmente conduzca a la ruptura de tales enlaces. Se tomaron en cuenta los estados más bajos en simetría C_s del átomo metálico: X (2º estado excitado, 1º estado excitado y estado fundamental) + YH_4 → H_3XYH → H + XYH_3 y XH + YH_3, con X = Au, Zn, Cd, Hg, Al, G, y con Y = C, Si, Ge, donde X denota al átomo metálico y, Y denota algún elemento del grupo IV. La relevancia se consideró en (a) el papel que desempeña la ocupación de las capas d, s, o p del átomo de metal en las interacciones con una molécula de metano o silano o germano, (b) el papel de cualquiera de los estados excitados singulete o doblete de metales en las barreras de potencial, y (c) el papel de las probabilidades de transición para diferentes familias de metales que reaccionan con estos gases, utilizando el ángulo H – X – Y como una coordenada de reacción. La ruptura del enlace Y – H de YH_4 es útil en la producción de películas amorfas hidrogenadas, necesarias en varios campos de la industria.

Interacción molécula-molécula

Consideremos la interacción Pt_4 + CO entre una superficie FCC de cuatro átomos de platino y una molécula de monóxido de carbono como caso intermolecular. La interacción del tetrámero de Pt_4 con la molécula de CO se estudia por medio del método de Hartree-Fock de campo autoconsistente (HF-SCF), usando potenciales relativistas de núcleo efectivo (RECP). Estos cálculos se continuaron con cálculos CI extensivos para mejorar la energía. Se consideraron los tres estados electrónicos más bajos X $^3A'$, a $^1A'$, b $^1A'$ del agregado completo para estudiar esta interacción. En la interacción Pt_4-CO de simetría C_{3v} frontal, el átomo de carbono de la molécula de CO está directamente frente a un átomo de platino del agregado piramidal de Pt_4. Este agregado se encuentra en sus tres estados electrónicos más bajos que pueden capturar la molécula de CO espontáneamente. Mientras que en la interacción lateral entre Pt_4 y CO en simetría C_{3v}, donde el átomo de carbono de la molécula de monóxido de carbono CO está directamente frente al centro de una cara formada por tres átomos de platino, el clúster de Pt_4 en sus dos estados electrónicos más bajos X $^3A'$ y a $^1A'$ puede capturar la molécula de CO después de superar una barrera de energía. Finalmente, estos resultados están en acuerdo con los datos experimentales de la interacción observada entre la molécula de CO

y la superficie de platino (ver figura 30). Se abunda con claridad en la discusión y en la estructura electrónica de este sistema en un artículo publicado por Castillo et al. (2005).

Figura 30. *Curvas de energía potencial C_{3v}: Pt_4 (x $^3A'$, a $^1A'$ and b $^1A'$) + CO*



Fuente: elaboración propia.

Capítulo 2. Métodos de primeros principios

Introducción

La dificultad fundamental al resolver de manera exacta la ecuación de onda para átomos con muchos electrones es que cada electrón repele a todos los demás electrones con una fuerza que va como la ley del inverso al cuadrado, así que el movimiento de cualquier electrón depende del movimiento de todos los demás. Si supusiéramos que al tratar con uno de los electrones podemos descartar las interacciones instantáneas con todos los demás electrones, y que ese electrón se mueve en un campo eléctrico que se obtiene de un promedio adecuado de todas las posiciones de todos los demás electrones, entonces cada electrón se describiría por una ecuación de onda que involucra solo sus coordenadas, y no las coordenadas de cualquier otro electrón; el efecto de todos los electrones se mostraría por la existencia de una función de energía potencial para el primer electrón.

Suponiendo que hay N electrones en un átomo, se propone una función de onda adecuada para cada uno de estos electrones. Se escoge a uno de los electrones y se encuentra el campo promedio de todos los demás electrones. Este campo promedio es simplemente el campo que se tendría si cada uno de los electrones fuera una nube-cargada de probabilidad. Si es necesario se promedia este campo sobre todos los ángulos para hacerlo esféricamente simétrico. Este proceso permite escribir, y luego resolver, la ecuación de onda para nuestro electrón escogido. Se obtiene una función de onda mejorada para ese electrón. Esta nueva función puede utilizarse para calcular el campo promedio para un segundo electrón, y nos permite tener una función mejorada para este electrón. El proceso continúa hasta tener un conjunto completo de orbitales mejorados. De la misma manera empezando con estas funciones se pueden mejorar una por una y en consecuencia se estarán calculando orbitales atómicos doblemente mejorados. Esta técnica se continúa hasta que después de varias iteraciones no se obtenga una diferencia apreciable entre ellas. Podemos decir que el conjunto de orbitales atómicos ha alcanzado la autoconsistencia; esto significa que, si escogemos cualquier electrón, su nube de carga es precisamente la que viene de resolver la ecuación de onda en la cual el campo de potencial es debido a toda la carga en el núcleo y la suma de las nubes de carga de todos los demás electrones. A este proceso se le conoce como método de campo autoconsistente (SCF, por sus siglas en inglés)

Método de campo autoconsistente

Las dos metodologías que han sido muy útiles para construir soluciones aproximadas a la ecuación de Schrödinger son el método variacional y la teoría de perturbaciones, los cuales forman la base para atacar muchos de los problemas de la estructura atómica y molecular.
La ecuación de Schrödinger multielectrónica no se puede separar en ecuaciones diferenciales ordinarias en términos de algún sistema de coordenadas conocido debido a los términos $\frac{1}{r_{ij}}$.
La no separabilidad de la ecuación de onda para sistemas de interés físico es lo más común, ya que solo pocos sistemas simples admiten soluciones exactas.

Considerando la ecuación de Schrödinger independiente del tiempo

$$H\psi = E\psi$$

donde ψ y E son la solución exacta y la energía del operador hamiltoniano, respectivamente. En general la función matemática del hamiltoniano será tal que la ecuación anterior no será separable en ningún sistema de coordenadas conocido. Entonces, no se puede esperar una expresión analítica de la función ψ.

Examinemos propiedades de lo que se conoce como funcional $\mathcal{E}(\xi)$ definido por

$$\mathcal{E}(\xi) = \frac{\langle \xi | H | \xi \rangle}{\langle \xi | \xi \rangle}$$

Donde H es el operador hamiltoniano del sistema y ξ es una función arbitraria, sujeta solamente a la restricción de que este es normalizable a 1, y cuya función exacta es ψ. Si ξ es idéntica a ψ, el funcional $\mathcal{E}$ es la energía

$$\mathcal{E}(\xi) = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\langle \psi | H | \psi \rangle}{1} = \langle \psi | E\psi \rangle = E\langle \psi | \psi \rangle = E$$

Consideremos una función ξ como una aproximación de prueba a la función exacta, y supongamos que ξ difiere de ψ no más allá de una variación de primer orden.

$$\xi = \psi + \delta\psi$$

calculemos ahora $(H - E)\xi$, esto es

$$\begin{aligned}(H - E)\xi &= (H - E)(\psi + \delta\psi) \\ &= H(\psi + \delta\psi) - E(\psi + \delta\psi) \\ &= H\psi + H\delta\psi - E\psi - E\delta\psi \\ &= H\delta\psi - E\delta\psi\end{aligned}$$

$$\therefore \quad (H - E)\xi = (H - E)\delta\psi$$

Es decir, se aniquiló la función exacta ψ. Calculemos el valor esperado de este último operador

$$\frac{\langle \xi|H-E|\xi\rangle}{\langle \xi|\xi\rangle} = \frac{\langle \xi|H-E|\delta\psi\rangle}{\langle \xi|\xi\rangle}$$

Dado que

$$\begin{aligned}\langle \xi|H-E|\delta\psi\rangle &= \langle \xi|H|\delta\psi\rangle - \langle \xi|E|\delta\psi\rangle \\ &= \langle H\xi|\delta\psi\rangle - \langle E\xi|\delta\psi\rangle \\ &= \langle (H-E)\xi|\delta\psi\rangle \\ &= \langle (H-E)\delta\psi|\delta\psi\rangle \\ &= \langle \delta\psi|H-E|\delta\psi\rangle\end{aligned}$$

y que

$$\begin{aligned}\langle \xi|H-E|\xi\rangle &= \langle \xi|H|\xi\rangle - \langle \xi|E|\xi\rangle \\ &= \langle H\xi|\xi\rangle - \langle E\xi|\xi\rangle \\ &= \langle (H-E)\xi|\xi\rangle \\ &= \langle (H-E)\delta\psi|\xi\rangle \\ &= \langle \delta\psi|H-E|\xi\rangle \\ &= \langle \delta\psi|(H-E)\xi\rangle \\ &= \langle \delta\psi|(H-E)\delta\psi\rangle \\ &= \langle \delta\psi|H-E|\delta\psi\rangle\end{aligned}$$

Entonces

$$\frac{\langle \xi|H-E|\xi\rangle}{\langle \xi|\xi\rangle} = \frac{\langle \delta\psi|H-E|\delta\psi\rangle}{\langle \xi|\xi\rangle}$$

$$\frac{\langle \xi|H|\xi\rangle - \langle \xi|E|\xi\rangle}{\langle \xi|\xi\rangle} = \frac{\langle \delta\psi|H-E|\delta\psi\rangle}{\langle \xi|\xi\rangle}$$

$$\frac{\langle \xi|H|\xi\rangle}{\langle \xi|\xi\rangle} - E\frac{\langle \xi|\xi\rangle}{\langle \xi|\xi\rangle} = \frac{\langle \delta\psi|H-E|\delta\psi\rangle}{\langle \xi|\xi\rangle}$$

Por lo tanto

$$\mathcal{E}(\xi) = E + \frac{\langle \delta\psi|H-E|\delta\psi\rangle}{\langle \xi|\xi\rangle}$$

Este resultado dice que el funcional $\mathcal{E}(\xi)$ es una aproximación exacta a E solo a segundo orden en $\delta\psi$. Ya que la función de onda de prueba ξ tiene un error a primer orden, mientras que la energía (total) tiene un error a segundo orden. Por otra parte, minimizar este funcional es equivalente a minimizar $J(\xi) = \langle \xi|H|\xi\rangle$ con la condición de que $\langle \xi|\xi\rangle = 1$. Entonces, en términos del multiplicador de Lagrange se tiene que

$$L(\xi) = J(\xi) - \lambda\langle\xi|\xi\rangle$$
$$L(\xi) = \langle\xi|H|\xi\rangle - \lambda\langle\xi|\xi\rangle$$

Ahora bien, queremos determinar la función de onda ξ que hace al funcional $L(\xi)$ obtener un valor estacionario (mínimo). La variación a primer orden del funcional $L(\xi)$ viene dada por [nota: $\delta L = L(\xi + \delta\xi) - L(\xi)$]:

$$L(\xi + \delta\xi) = J(\xi + \delta\xi) - \lambda\langle\xi + \delta\xi|\xi + \delta\xi\rangle$$
$$= \langle\xi + \delta\xi|H|\xi + \delta\xi\rangle - \lambda\langle\xi + \delta\xi|\xi + \delta\xi\rangle$$

Entonces

$$\delta L = \langle\xi + \delta\xi|H|\xi + \delta\xi\rangle - \lambda\langle\xi + \delta\xi|\xi + \delta\xi\rangle - \langle\xi|H|\xi\rangle + \lambda\langle\xi|\xi\rangle$$
$$= \langle\xi|H|\xi + \delta\xi\rangle + \langle\delta\xi|H|\xi + \delta\xi\rangle - \lambda\langle\xi|\xi + \delta\xi\rangle - \lambda\langle\delta\xi|\xi + \delta\xi\rangle - \langle\xi|H|\xi\rangle + \lambda\langle\xi|\xi\rangle$$
$$= \langle\xi|H|\xi\rangle + \langle\xi|H|\delta\xi\rangle + \langle\delta\xi|H|\xi\rangle + \langle\delta\xi|H|\delta\xi\rangle - \lambda\langle\xi|\xi\rangle - \lambda\langle\xi|\delta\xi\rangle - \lambda\langle\delta\xi|\xi\rangle - \lambda\langle\delta\xi|\delta\xi\rangle$$
$$- \langle\xi|H|\xi\rangle + \lambda\langle\xi|\xi\rangle$$
$$= \langle\xi|H|\delta\xi\rangle + \langle\delta\xi|H|\xi\rangle - \lambda\langle\xi|\delta\xi\rangle - \lambda\langle\delta\xi|\xi\rangle$$

Por lo tanto, la derivada variacional $\delta L(\xi) = \delta(\langle\xi|H|\xi\rangle - \lambda\langle\xi|\xi\rangle)$ está dada por:

$$\delta L(\xi) = \langle\xi|H|\delta\xi\rangle + \langle\delta\xi|H|\xi\rangle - \lambda\langle\xi|\delta\xi\rangle - \lambda\langle\delta\xi|\xi\rangle$$

Minimizar significa calcular $\delta L(\xi) = 0$

Entonces

$$\langle\xi|H|\delta\xi\rangle + \langle\delta\xi|H|\xi\rangle - \lambda\langle\xi|\delta\xi\rangle - \lambda\langle\delta\xi|\xi\rangle = 0$$
$$\langle\xi|H|\delta\xi\rangle - \lambda\langle\xi|\delta\xi\rangle + \langle\delta\xi|H|\xi\rangle - \lambda\langle\delta\xi|\xi\rangle = 0$$
$$\langle H\xi|\delta\xi\rangle - \langle\lambda\xi|\delta\xi\rangle + \langle\delta\xi|(H - \lambda)\xi\rangle = 0$$
$$\langle(H\xi - \lambda^*\xi)|\delta\xi\rangle + \langle\delta\xi|(H - \lambda)\xi\rangle = 0$$

Por lo tanto, por una parte $((H\xi)^* - \lambda^*\xi^*) = 0$ y, por otra parte:

$$(H - \lambda)\xi_{esp} = 0 \quad \Longrightarrow \quad H\xi_{esp} - \lambda\xi_{esp} = 0 \quad \therefore \quad H\xi_{esp} = \lambda\xi_{esp}$$

Esto es que cuando se minimiza la energía por un cálculo variacional, intrínsicamente se está resolviendo la ecuación de Schrödinger. Esta ecuación se satisface para cualquier variación arbitraria $\delta\xi$, siempre que

$$H\xi - \lambda\xi = 0$$

La cual es exactamente la ecuación de Schrödinger. La condición de que $\delta\langle H\rangle$ sea cero es suficiente para que ξ sea una función propia del hamiltoniano. Se sugiere que el multiplicador de Lagrange sea idéntico al funcional $\mathcal{E}(\xi)$.

Como sabemos el valor del funcional de energía para una función ξ, viene dado por

$$\mathcal{E}(\xi) = \frac{\langle \xi | H | \xi \rangle}{\langle \xi | \xi \rangle}$$

Consideremos ahora un conjunto completo de funciones propias ortonormales $\{\psi_i\}$ del hamiltoniano, y sean $\{E_i\}$ los valores propios correspondientes. Expresemos la función ξ en términos de esta base de funciones

$$\xi = \sum_{i=1}^{N} c_i \psi_i$$

Sustituyendo esta expresión en la anterior se tiene que

$$\mathcal{E}(\xi) = \frac{\left\langle \sum_{i=1}^{N} c_i \psi_i \middle| H \middle| \sum_{i=1}^{N} c_i \psi_i \right\rangle}{\left\langle \sum_{i=1}^{N} c_i \psi_i \middle| \sum_{i=1}^{N} c_i \psi_i \right\rangle}$$

y como $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ (p. ej.: $\langle \xi | \xi \rangle = 1 \; o \; bien \; \int \xi^* \xi dV = 1$) se tiene que

$$\mathcal{E}(\xi) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} c_i^* c_j \langle \psi_i | H | \psi_j \rangle}{\sum_{i=1}^{N} \sum_{j=1}^{N} c_i^* c_j \langle \psi_i | \psi_j \rangle}$$

$$\mathcal{E}(\xi) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} c_i^* c_j \langle \psi_i | H | \psi_j \rangle}{\sum_{i=1}^{N} |c_i|^2} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} c_i c_j \langle \psi_i | E_j | \psi_j \rangle}{\sum_{i=1}^{N} |c_i|^2}$$

debido a que $H\psi_j = E_j \psi_j$ es la ecuación de Schrödinger, entonces

$$\mathcal{E}(\xi) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} E_j c_i c_j \langle \psi_i | \psi_j \rangle}{\sum_{i=1}^{N} |c_i|^2}$$

$$\therefore \qquad \mathcal{E} = \frac{\sum_{j=1}^{N} |c_j|^2 E_j}{\sum_{i=1}^{N} |c_i|^2}$$

Nótese que $\left\langle \sum_{i=1}^{N} c_i \psi_i \middle| \sum_{i=1}^{N} c_i \psi_i \right\rangle = 1$ implica que

$$\sum_{i=1}^{N} \sum_{j=1}^{N} c_i^* c_j \langle \psi_i | \psi_j \rangle = 1 \qquad \therefore \qquad \sum_{i=1}^{N} |c_i|^2 = 1$$

Esto es

$$\mathcal{E} = \sum_{j=1}^{N} |c_j|^2 E_j$$

Ahora bien, restando a $\mathcal{E}$ el valor de la energía propia del estado base ($H\psi_0 = E_0\psi_0$), se tiene que

$$\mathcal{E} - E_0 = \sum_{j=1}^{N} |c_j|^2 (E_j - E_0)$$

Como $E_j - E_0 \geq 0$ entonces

$$\mathcal{E} - E_0 \geq 0$$

Este resultado se conoce como el teorema del límite superior de Eckart. *Este teorema muestra que cualquier función de prueba (la cual está normalizada) lleva a un valor de la energía que nunca es más baja que el valor de la energía del estado base.*

Dado que en la práctica no es posible expandir la función de onda de prueba en términos de un número infinito de funciones de onda de una base completa de funciones, entonces una expansión en términos de un número finito de tales funciones podría proporcionar una aproximación razonable. A tal conjunto finito se le conoce como una base truncada. En la práctica uno busca un número finito de *n* funciones, intentando encontrar aquellas *n* funciones particulares, las cuales llevan a la energía mínima, para un sistema dado, algunos conjuntos de funciones bases pueden ser más eficientes que otros, en el sentido en que pocas funciones base son requeridas para llevar a una energía dada.

Ahora, expandiendo la función de prueba ξ_k a una base arbitraria $\{\phi\}$ y truncando la expansión en la *n*-ésima función, se tiene

$$\xi_k = \sum_{i=1}^{n} \phi_i c_{ik}$$

aquí suponemos que las funciones $\{\phi_j\}$ son independientes, pero no necesariamente ortonormales.

Definimos las matrices

$$(1 \times n) \quad \longrightarrow \quad \phi = (\phi_1 \phi_2 \phi_3 \cdots \phi_N)$$

$$(n \times 1) \quad \longrightarrow \quad \phi^T = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_N \end{pmatrix}$$

$$\phi\phi^T = (\phi_1\phi_2\phi_3 \cdots \phi_N)\begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_N \end{pmatrix}$$

Además, definamos la matriz

$$c_k = \begin{pmatrix} c_{1k} \\ c_{2k} \\ c_{3k} \\ \vdots \\ c_{Nk} \end{pmatrix} \qquad\qquad c_k^\dagger = (c_{1k}^* c_{2k}^* c_{3k}^* \cdots c_{Nk}^*)$$

entonces, podemos volver a definir ξ_k como

$$\xi_k = \phi c_k$$

y definiendo ahora $\phi^\dagger = \begin{pmatrix} \phi_1^* \\ \phi_2^* \\ \phi_3^* \\ \vdots \\ \phi_N^* \end{pmatrix}$, $\quad (\mathbb{H})_{ij} = \langle \phi_i | H | \phi_j \rangle$, $\quad$ y $\quad (\Delta)_{ij} = \langle \phi_i | \phi_j \rangle$

ahora podemos volver a expresar el funcional de la energía como

$$\mathcal{E} = \frac{c_k^\dagger \mathbb{H} c_k}{c_k^\dagger \, \Delta \, c_k}.$$

Aplicando una variación a esta ecuación se tiene que

$$\delta\mathcal{E} = \delta\left(\frac{c_k^\dagger \mathbb{H} c_k}{c_k^\dagger \, \Delta \, c_k}\right)$$

$$\delta\mathcal{E} = \frac{c_k^\dagger \, \Delta \, c_k \delta\left(c_k^\dagger \mathbb{H} c_k\right) - c_k^\dagger \mathbb{H} c_k \delta\left(c_k^\dagger \, \Delta \, c_k\right)}{\left(c_k^\dagger \, \Delta \, c_k\right)^2}$$

Nótese que $\delta c_k = \begin{pmatrix} \delta c_{1k} \\ \delta c_{2k} \\ \delta c_{3k} \\ \vdots \\ \delta c_{Nk} \end{pmatrix}$

$$\delta\mathcal{E} = \frac{c_k^\dagger \vartriangle c_k\{\delta c_k^\dagger \mathbb{H} c_k + c_k^\dagger \mathbb{H}\delta c_k\} - c_k^\dagger \mathbb{H} c_k\{\delta c_k^\dagger \vartriangle c_k + c_k^\dagger \vartriangle \delta c_k\}}{\left(c_k^\dagger \vartriangle c_k\right)^2}$$

$$\delta\mathcal{E} = \frac{\delta c_k^\dagger \mathbb{H} c_k}{c_k^\dagger \vartriangle c_k} + \frac{c_k^* \mathbb{H}\delta c_k}{c_k^\dagger \vartriangle c_k} - \frac{1}{c_k^\dagger \vartriangle c_k}\mathcal{E}\delta c_k^\dagger \vartriangle c_k - \frac{1}{c_k^\dagger \vartriangle c_k}\mathcal{E} c_k^* \vartriangle \delta c_k$$

Como estamos minimizando $\delta\mathcal{E} = 0$, entonces

$$\frac{1}{c_k^\dagger \vartriangle c_k}\{\delta c_k^\dagger \mathbb{H} c_k - \mathcal{E}\delta c_k^\dagger \vartriangle c_k\} + el\ complejo\ conjugado = 0$$

$$\delta c_k^\dagger\{\mathbb{H} c_k - \mathcal{E} \vartriangle c_k\} = 0$$

Como $\delta c_k^\dagger \neq 0$, se tiene que

$$(\mathbb{H} c_k - \mathcal{E} \vartriangle c_k) = 0$$

$$\therefore \quad (\mathbb{H} - \mathcal{E} \vartriangle) c_k = 0$$

La solución no trivial de esta ecuación requiere de *n* raíces del determinante secular

$$det(\mathbb{H} - \mathcal{E} \vartriangle) = 0$$

etiquetando las n raíces $\epsilon_1, \epsilon_2, \dots, \epsilon_N$ que corresponden a los valores propios $c_1, c_2, \dots, c_N$, respectivamente. Ahora definimos las matrices

$$C = (C_1 C_2 C_3 \cdots C_N)$$

Con $C_1 = \begin{pmatrix} c_{11} \\ c_{21} \\ c_{31} \\ \vdots \\ c_{N1} \end{pmatrix}$, $C_2 = \begin{pmatrix} c_{12} \\ c_{22} \\ c_{32} \\ \vdots \\ c_{N2} \end{pmatrix}$, $\cdots$, se tiene que

$$(\mathcal{E})_{ij} = C_i \delta_{ij}$$

y escribiendo la ecuación de Schrödinger en forma matricial

$$\mathbb{H} C = \vartriangle C\mathcal{E}$$

En la eventualidad de que la base $\{\phi_i\}$ sea ortogonal, se tiene que

$$\mathcal{H} C = C\mathcal{E}$$

Orbitales espaciales y espín-orbitales

Se define como orbital electrónico a la función de onda que describe la distribución espacial de un electrón en un sistema molecular. La ecuación de Schrödinger no relativista no

considera los términos que involucran al espín electrónico, por lo que se hace necesario incluir esta propiedad electrónica en el planteamiento de la solución.

En general un orbital electrónico dependerá tanto de las coordenadas espaciales como del espín. Una base completa para describir el espín de un electrón consiste en dos funciones ortonormales $\alpha(\omega)$ y $\beta(\omega)$. Se define como espín orbital a la función de onda de un electrón que describe tanto a su distribución espacial como a su espín.

Con cada orbital $\varphi(\boldsymbol{r})$ se pueden construir dos espín-orbitales, uno correspondiente al espín con proyección $m_S = 1$ y el otro con $m_S = -1$, multiplicando a este orbital por las funciones $\alpha(\omega)$ y $\beta(\omega)$, respectivamente. Entonces, un conjunto completo que describe al espín de un electrón consiste en dos funciones ortonormales $\alpha(\omega)$ y $\beta(\omega)$, *i.e.*, espín hacia arriba $(\uparrow)$ y espín hacia abajo $(\downarrow)$. La función de onda de un electrón que describe tanto su distribución espacial como su espín es un espín-orbital, $\chi(\boldsymbol{x})$, donde $\boldsymbol{x}$ indica las coordenadas de espacio y de espín. De cada orbital espacial $\varphi(\boldsymbol{r})$, uno puede formar dos espín-orbitales

$$\chi(\boldsymbol{x}) = \begin{cases} \varphi(\boldsymbol{r})\alpha(\omega) \\ \varphi(\boldsymbol{r})\beta(\omega) \end{cases}$$

Nótese que:

$$\alpha(\omega)\beta(\omega) = 0$$
$$\alpha^*(\omega)\alpha(\omega) = 1$$
$$\beta^*(\omega)\beta(\omega) = 1$$

Consideremos un sistema de *N* electrones cuyas coordenadas vienen dadas por $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N$, El determinante de Slater construido con un conjunto de N funciones espín-orbitales $\{\chi_1(\boldsymbol{x}_1), \chi_2(\boldsymbol{x}_2), \cdots, \chi_N(\boldsymbol{x}_N)\}$ viene dado por

$$\psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N) = \frac{1}{(N!)^{\frac{1}{2}}} \begin{vmatrix} \chi_1(\boldsymbol{x}_1) & \chi_1(\boldsymbol{x}_2) & \chi_1(\boldsymbol{x}_3) & \dots & \chi_1(\boldsymbol{x}_N) \\ \chi_2(\boldsymbol{x}_1) & \chi_2(\boldsymbol{x}_2) & \chi_2(\boldsymbol{x}_3) & \dots & \chi_2(\boldsymbol{x}_N) \\ \chi_3(\boldsymbol{x}_1) & \chi_3(\boldsymbol{x}_2) & \chi_3(\boldsymbol{x}_3) & \dots & \chi_3(\boldsymbol{x}_N) \\ \vdots & \vdots & \vdots & & \vdots \\ \chi_N(\boldsymbol{x}_1) & \chi_N(\boldsymbol{x}_2) & \chi_N(\boldsymbol{x}_3) & \dots & \chi_N(\boldsymbol{x}_N) \end{vmatrix}$$

Estas funciones se representan como

$$\psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N) = (N!)^{-\frac{1}{2}} |\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2) \cdots \chi_N(\boldsymbol{x}_N)|$$

Donde $\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2) \cdots \chi_N(\boldsymbol{x}_N)$ es la traza del determinante en ambas expresiones.

El determinante de Slater es una función de onda que está constituida por la suma de los productos de N funciones χ_i, esto es $\chi_1\chi_2\cdots\chi_N$. Ejemplo

$$\begin{vmatrix}\chi_1(1) & \chi_2(1)\\ \chi_1(2) & \chi_2(2)\end{vmatrix} = \begin{vmatrix}\chi_1(1) & \chi_1(2)\\ \chi_2(1) & \chi_2(2)\end{vmatrix} = \chi_1(1)\chi_2(2) - \chi_1(2)\chi_2(1)$$

Tiene 2! términos. Un determinante de 3×3 tiene 3! términos en la solución

$$\begin{vmatrix}\chi_1(1) & \chi_2(1) & \chi_3(1)\\ \chi_1(2) & \chi_2(2) & \chi_3(2)\\ \chi_1(3) & \chi_2(3) & \chi_3(3)\end{vmatrix} = \begin{cases}\chi_1(1)\chi_2(2)\chi_3(3) + \chi_2(1)\chi_3(2)\chi_1(3) + \chi_3(1)\chi_1(2)\chi_2(3)\\ -\chi_3(1)\chi_2(2)\chi_1(3) - \chi_1(1)\chi_3(2)\chi_2(3) - \chi_2(1)\chi_1(2)\chi_3(3)\end{cases}$$

También se pueden usar permutaciones. Las permutaciones van como $(-1)^P$, si la permutación es par el signo es +, y si la permutación es impar el signo es −. Los subíndices de las funciones son

$$\begin{matrix}1 & 2 & 3\\ 3 & 2 & 1\end{matrix} \qquad \begin{matrix}2 & 3 & 1\\ 1 & 3 & 2\end{matrix} \qquad \begin{matrix}3 & 1 & 2\\ 2 & 1 & 3\end{matrix}$$

Tres permutaciones en la línea de abajo del primer agrupamiento de izquierda a derecha, dos permutaciones en la línea de arriba y una en la línea de abajo del segundo agrupamiento, dos permutaciones en la línea de arriba y una en la línea de abajo del tercer agrupamiento. En total son tres permutaciones en cada caso.

Entonces, también podemos fijar los subíndices de la función de onda, en vez de fijar los índices de coordenada. Esto se ve si resolvemos el determinante por cofactores, como se hace comúnmente, es decir

$$\begin{vmatrix}\chi_1(1) & \chi_2(1) & \chi_3(1)\\ \chi_1(2) & \chi_2(2) & \chi_3(2)\\ \chi_1(3) & \chi_2(3) & \chi_3(3)\end{vmatrix} = \begin{cases}\chi_1(1)\chi_2(2)\chi_3(3) - \chi_1(1)\chi_2(3)\chi_3(2)\\ -\chi_1(2)\chi_2(1)\chi_3(3) + \chi_1(3)\chi_2(1)\chi_3(2)\\ \chi_1(2)\chi_2(3)\chi_3(1) - \chi_1(3)\chi_2(2)\chi_3(1)\end{cases}$$

En donde cada producto se identifica con una permutación de las coordenadas de los electrones llevada a cabo en este producto de funciones, o equivalentemente el determinante de Slater está construido por la suma de los productos de las N funciones χ_i, en donde cada producto corresponde a una permutación de los espín-orbitales manteniendo fijos los argumentos $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$.

La función determinante de Slater se puede escribir en forma simbólica como

$$det\phi = \mathcal{O}\phi$$

Donde $\mathcal{O}$ es un operador simetrizador. Para una función de onda *N*-electrónica en la aproximación orbital tiene la forma general

$$\begin{vmatrix} \phi_1(1) & \phi_1(2) & \phi_1(3) \\ \phi_2(1) & \phi_2(2) & \phi_2(3) \\ \phi_3(1) & \phi_3(2) & \phi_3(3) \end{vmatrix} = (3!)^{-\frac{1}{2}} \sum_P (-1)^P \mathbb{P}\phi_1(1)\phi_2(2)\phi_3(3)$$

Donde $\mathbb{P}$ es un operador de permutación. Por ejemplo, el intercambio $3 \rightarrow 2$ de $\mathbb{P} \Longrightarrow \phi_1(1)\phi_2(2)\phi_3(3)$ deja $(-1)^1\phi_1(1)\phi_2(3)\phi_3(2)$. Por lo tanto

$$(N!)^{-\frac{1}{2}}|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)| = (N!)^{-\frac{1}{2}} \sum_P (-1)^P \mathbb{P}\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)$$

Para una función de onda N-electrónica en la aproximación orbital, este operador orbital o bien operador antisimetrizador de funciones de onda de N partículas es

$$\mathcal{O} = (N!)^{-\frac{1}{2}} \sum_P (-1)^P \mathbb{P}$$

En el presente contexto el operador lleva a cabo $N!$ diferentes permutaciones de las coordenadas de los electrones en el producto de las N funciones para formar los componentes de una función antisimétrica. La parte de la permutación al antisimetrizador puede ser escrita en la forma expandida

$$\sum_P (-1)^P \mathbb{P} = 1 - \sum_{ij} \mathbb{P}_{ij} + \sum_{ijk} \mathbb{P}_{ijk}$$

Tomemos como ejemplo cuatro partículas $\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4)$. Aplicando el operador se tiene en un ejemplo simple:

$$\mathbb{P}_{21}\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) = \phi_1(2)\phi_2(1)\phi_3(3)\phi_4(4)$$

Más explícitamente

$$\mathcal{O}\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) = \frac{1}{(4!)^{\frac{1}{2}}}\{1 + \mathbb{P}_{21} - \mathbb{P}_{43} + \mathbb{P}_{41} - \mathbb{P}_{32} + \mathbb{P}_{31} - \mathbb{P}_{42}$$

$$+\mathbb{P}_{423} + \mathbb{P}_{342} - \mathbb{P}_{341} - \mathbb{P}_{413} + \mathbb{P}_{241} - \mathbb{P}_{412} + \mathbb{P}_{231} + \mathbb{P}_{312} - \mathbb{P}_{2341} - \mathbb{P}_{4321}$$

$$-\mathbb{P}_{3421} + \mathbb{P}_{4312} - \mathbb{P}_{3142} - \mathbb{P}_{2143} + \mathbb{P}_{412} - \mathbb{P}_{2413} + \mathbb{P}_{3412}\Big\}\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4)$$

$$\mathcal{O}\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) = \frac{1}{\sqrt{4!}}\{\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) - \phi_1(2)\phi_2(1)\phi_3(3)\phi_4(4)$$

$$-\phi_1(1)\phi_2(2)\phi_3(4)\phi_4(3) - \phi_1(4)\phi_2(2)\phi_3(3)\phi_4(1) - \phi_1(1)\phi_2(3)\phi_3(2)\phi_4(4)$$
$$-\phi_1(3)\phi_2(2)\phi_3(1)\phi_4(4) - \phi_1(1)\phi_2(4)\phi_3(3)\phi_4(2) + \phi_1(1)\phi_2(4)\phi_3(2)\phi_4(3)$$
$$+\phi_1(1)\phi_2(3)\phi_3(4)\phi_4(2) + \phi_1(3)\phi_2(2)\phi_3(4)\phi_4(1) + \phi_1(4)\phi_2(1)\phi_3(3)\phi_4(2)$$
$$+\phi_1(2)\phi_2(4)\phi_3(3)\phi_4(1) + \phi_1(4)\phi_2(2)\phi_3(1)\phi_4(3) + \phi_1(2)\phi_2(3)\phi_3(1)\phi_4(4)$$
$$+\phi_1(3)\phi_2(1)\phi_3(2)\phi_4(4) - \phi_1(2)\phi_2(3)\phi_3(4)\phi_4(1) + \phi_1(4)\phi_2(3)\phi_3(2)\phi_4(1)$$
$$-\phi_1(3)\phi_2(4)\phi_3(2)\phi_4(1) - \phi_1(4)\phi_2(3)\phi_3(1)\phi_4(2) - \phi_1(3)\phi_2(1)\phi_3(4)\phi_4(2)$$
$$+\phi_1(2)\phi_2(1)\phi_3(4)\phi_4(3) - \phi_1(4)\phi_2(1)\phi_3(2)\phi_4(3) - \phi_1(2)\phi_2(4)\phi_3(1)\phi_4(3)$$
$$+\phi_1(3)\phi_2(4)\phi_3(1)\phi_4(2)\}$$

En general

$$\mathcal{O} = (N!)^{-\frac{1}{2}} \sum_P (-1)^P \mathbb{P} = \frac{1}{(N!)^{\frac{1}{2}}} \left\{ 1 - \sum_{ij} \mathbb{P}_{ij} + \sum_{ijk} \mathbb{P}_{ijk} - \cdots \right\}$$

donde $\mathbb{P}_{ijk}$ es el operador que lleva a cabo todas las posibles permutaciones de tres partículas ijk. Con el propósito de hacer un mejor uso de las ventajas de la notación de la antisimetrización, es necesario presentar algunas de sus propiedades. Consideremos una función de onda N-electrónica dada por

$$\phi = \phi(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N)$$

La operación de $\mathbb{P}$ es permutar las coordenadas electrónicas a un nuevo orden $\boldsymbol{x}'_1, \boldsymbol{x}'_2, \boldsymbol{x}'_3, \dots, \boldsymbol{x}'_N$ donde cada $\boldsymbol{x}'_i$ es una de las coordenadas. Designamos a esta permutación simbólica por

$$\mathbb{P}\phi(\tau) = \phi(\tau')$$

donde τ representa la colección de coordenadas en forma no permutada y τ' la colección de coordenadas en el orden permutado. Ahora en términos de la permutación inversa esta relación viene dada por

$$\phi(\tau) = \mathbb{P}^{-1}\phi(\tau')$$

Ahora consideremos la integral definida

$$\langle \mathbb{P}\phi_1(\tau) | \phi_2(\tau) \rangle \text{ equivalente a } \int \left(\mathbb{P}\phi_1(\tau)\right)^* \phi_2(\tau)\, d\boldsymbol{x}_1 d\boldsymbol{x}_2 \dots d\boldsymbol{x}_N$$

Recordemos que el operador daga (†) tiene la propiedad $\langle \hat{A}\phi | \varphi \rangle = \langle \phi | A^\dagger \varphi \rangle$. Con base en la definición del operador adjunto $\mathbb{P}^\dagger$ se tiene que

$$\langle \mathbb{P}\phi_1(\tau)|\phi_2(\tau)\rangle = \langle \phi_1(\tau')|\mathbb{P}^\dagger\phi_2(\tau)\rangle$$
$$= \langle \phi_1(\tau')|\phi_2(\tau)\rangle$$
$$= \langle \phi_1(\tau')|\mathbb{P}^{-1}\phi_2(\tau')\rangle$$

Ahora, al renombrar coordenadas:

$$\mathbb{P}^\dagger = \mathbb{P}^{-1}$$

Consideremos el adjunto del operador antisimetrizador

$$\mathcal{O}^\dagger = (N!)^{-\frac{1}{2}}\sum_{P^\dagger}(-1)^{P^\dagger}\mathbb{P}^\dagger$$
$$= (N!)^{-\frac{1}{2}}\sum_{P^{-1}}(-1)^{P}\mathbb{P}^{-1}$$
$$= (N!)^{-\frac{1}{2}}\sum_{P}(-1)^{P}\mathbb{P}^{-1}$$

para permutaciones de partículas. Por lo tanto, se ha demostrado que

$$\mathcal{O} = \mathcal{O}^\dagger$$

Tomemos en cuenta ahora el operador $\mathcal{O}^2$, se tiene que

$$\mathcal{O}^2 = \mathcal{O}\mathcal{O}$$
$$= \left\{(N!)^{-\frac{1}{2}}\sum_{P_j}(-1)^{P_j}\mathbb{P}_j\right\}\left\{(N!)^{-\frac{1}{2}}\sum_{P_j}(-1)^{P_j}\mathbb{P}_j\right\}$$
$$= (N!)^{-1}\sum_{P_j}(-1)^{P_j}\mathbb{P}_j\{(-1)^{P_1}\mathbb{P}_1 + (-1)^{P_2}\mathbb{P}_2 + \cdots + (-1)^{P_k}\mathbb{P}_k\}$$
$$= (N!)^{-1}\left\{\sum_{P_j}(-1)^{P_j}(-1)^{P_1}\mathbb{P}_j\,\mathbb{P}_1 + \sum_{P_j}(-1)^{P_j+P_2}\mathbb{P}_j\,\mathbb{P}_2 + \cdots + \sum_{P_j}(-1)^{P_j+P_k}\mathbb{P}_j\,\mathbb{P}_k\right\}$$

Como

$$\sum_{P_j}(-1)^{P_j+P_k}\mathbb{P}_j\,\mathbb{P}_k = \sum_{P'}(-1)^{P'}\mathbb{P}' = (N!)^{\frac{1}{2}}(N!)^{-\frac{1}{2}}\sum_{P'}(-1)^{P'}\mathbb{P}' = (N!)^{\frac{1}{2}}\mathcal{O}$$

$$\mathcal{O}^2 = (N!)^{-1}\left(N!\sum_{P_j}(-1)^{P_j+P_k}\mathbb{P}_j\,\mathbb{P}_k\right)$$

$$\mathcal{O}^2 = (N!)^{-1}\left(N!\,(N!)^{\frac{1}{2}}\mathcal{O}\right)$$

$$\therefore \qquad \mathcal{O}^2 = (N!)^{\frac{1}{2}}\mathcal{O}$$

Principio de exclusión de Pauli y determinante de Slater

El principio de Pauli en forma débil dice que *en un átomo multielectrónico nunca podrán coexistir más de un electrón en el mismo estado cuántico*, y en forma fuerte dice que *la función de onda total debe ser antisimétrica con respecto al intercambio de cualquier par de electrones.*

Si se quiere considerar una función de onda para una colección de electrones, *i.e.*, funciones de onda para *N*-electrones, tenemos que el hamiltoniano es una suma de hamiltonianos de un electrón. La función de onda que es una eigenfunción de este hamiltoniano es un simple producto de funciones de onda espín-orbitales para cada electrón que está dada por el producto de Hartree.

$$\Psi^{HP}(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N) = \chi_i(\boldsymbol{x}_1)\chi_j(\boldsymbol{x}_2)\cdots\chi_k(\boldsymbol{x}_N)$$

El producto de Hartree no satisface este principio de antisimetría. Sin embargo, se pueden obtener funciones de onda correctamente antisimetrizadas. Considerando el caso de dos electrones en el que se ocupan los espín-orbitales χ_i y χ_j, poniendo el electrón uno en χ_i y el electrón dos en χ_j, tenemos

$$\Psi_{12}^{HP}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \chi_i(\boldsymbol{x}_1)\chi_j(\boldsymbol{x}_2)$$

O bien, si ponemos el electrón uno en χ_j y el electrón dos en χ_i, tenemos

$$\Psi_{21}^{HP}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \chi_i(\boldsymbol{x}_2)\chi_j(\boldsymbol{x}_1)$$

Estos productos de Hartree distinguen entre electrones, sin embargo, se puede obtener una función de onda que no lo haga, y que satisfaga los requerimientos del principio de antisimetría tomando la combinación lineal apropiada con estos productos de Hartree.

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = 2^{-1/2}\left(\chi_i(\boldsymbol{x}_1)\chi_j(\boldsymbol{x}_2) - \chi_j(\boldsymbol{x}_1)\chi_i(\boldsymbol{x}_2)\right)$$

El factor $2^{-1/2}$ es un factor de normalización. El signo menos asegura que $\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ es antisimétrica con respecto al intercambio de las coordenadas de electrones uno y dos.

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = -\Psi(\boldsymbol{x}_2, \boldsymbol{x}_1)$$

Se puede reescribir conforme al determinante de Slater como

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = 2^{-1/2}\begin{vmatrix}\chi_i(\boldsymbol{x}_1) & \chi_j(\boldsymbol{x}_1)\\ \chi_i(\boldsymbol{x}_2) & \chi_j(\boldsymbol{x}_2)\end{vmatrix}$$

En notación más corta este determinante se puede escribir como

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N) = |\chi_i(\boldsymbol{x}_1)\chi_j(\boldsymbol{x}_2)\cdots\chi_k(\boldsymbol{x}_N)\rangle$$

El determinante de Slater incorpora correlación de intercambio, lo que significa que el movimiento de dos electrones con espines paralelos está correlacionado. El movimiento de los electrones con espines opuestos no lo está.

Ecuaciones de Hartree-Fock

Anteriormente se presentó la conveniencia de expresar la función de onda que describe un estado electrónico de un sistema molecular como un determinante de Slater o una suma de ellos, generados por un conjunto de espín orbitales (χ_c) a determinar mediante la ecuación de Schrödinger en forma directa en cosas muy simples, por lo que se debe recurrir a procedimientos no directos pero equivalentes: el método variacional, el cual consiste en determinar un conjunto de funciones φ_i para las cuales el valor esperado *E* de la energía de un estado molecular es mínimo

$$E = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$$

Las repulsiones electrón-electrón son fundamentales y se deben incluir en cualquier método preciso. En el método de Hartree-Fock (HF) se busca una función de onda con las repulsiones electrón-electrón tratadas en forma promedio. Se considera a cada electrón dentro del campo de los núcleos y del campo promedio de los N-1 electrones restantes. La función de onda antisimétrica más simple para representar al estado fundamental de una molécula o de un sistema molecular con *N* electrones, es un solo determinante de Slater construido con N espín orbitales.

$$\psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N) = (N!)^{-\frac{1}{2}}|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)|$$

Con esto se toma el valor esperado de la energía *E*. Nótese que

$$\begin{aligned}(N!)^{-\frac{1}{2}}|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)| &= \mathcal{O}\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\\ &= \mathcal{O}\xi(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \dots, \boldsymbol{x}_N)\end{aligned}$$

entonces

$$\langle\psi|\psi\rangle = \langle\mathcal{O}\xi|\mathcal{O}\xi\rangle = \langle\xi|\mathcal{O}^\dagger\mathcal{O}\xi\rangle$$

Como $\mathcal{O}^\dagger = \mathcal{O}$

$$\langle\psi|\psi\rangle = \langle\xi|\mathcal{O}^2\xi\rangle$$

$$\text{y} \quad \mathcal{O}^2 = (N!)^{\frac{1}{2}}\mathcal{O}$$

$$\Longrightarrow \quad \langle\psi|\psi\rangle = \left\langle\xi\middle|(N!)^{\frac{1}{2}}\mathcal{O}\xi\right\rangle$$

$$= (N!)^{\frac{1}{2}}\langle\xi|\mathcal{O}\xi\rangle$$

$$= (N!)^{\frac{1}{2}}\left\langle\xi\middle|(N!)^{-\frac{1}{2}}\textstyle\sum_{P_j}(-1)^{P_j}\mathbb{P}_j\,\xi\right\rangle$$

$$= \left\langle\xi\middle|\textstyle\sum_{P_j}(-1)^{P_j}\mathbb{P}_j\,\xi\right\rangle$$

$$= \langle\xi|\big(1 - \textstyle\sum_{ij}\mathbb{P}_{ij} + \sum_{ijk}\mathbb{P}_{ijk} - \cdots\big)\xi\rangle$$

$$= \langle\xi|\xi\rangle - \langle\xi|\textstyle\sum_{ij}\mathbb{P}_{ij}\,\xi\rangle + \langle\xi|\sum_{ijk}\mathbb{P}_{ijk}\,\xi\rangle - \cdots$$

tomando un conjunto de funciones orbitales φ_i ortonormales, esto es $\langle\varphi_i|\varphi_i\rangle = \delta_{ij}$

Como estamos considerando capa cerrada, entonces

$$\psi = (N!)^{-\frac{1}{2}}|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)| = (N!)^{-\frac{1}{2}}\mathcal{O}\xi$$

$$= (N!)^{-\frac{1}{2}}\mathcal{O}\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)$$

Tomando $\langle\xi|\xi\rangle$ entonces

$$\langle\xi|\xi\rangle = \int \chi_1^*(\boldsymbol{x}_1)\chi_2^*(\boldsymbol{x}_2)\chi_3^*(\boldsymbol{x}_3)\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\chi_3(\boldsymbol{x}_3)\,d\boldsymbol{x}_1 d\boldsymbol{x}_2 d\boldsymbol{x}_3 \ldots$$

$$= \int \chi_1^*(\boldsymbol{x}_1)\chi_1(\boldsymbol{x}_1)d\boldsymbol{x}_1 \int \chi_2^*(\boldsymbol{x}_2)\chi_2(\boldsymbol{x}_2)d\boldsymbol{x}_2 \int \chi_3^*(\boldsymbol{x}_3)\chi_3(\boldsymbol{x}_3)\,d\boldsymbol{x}_3 \ldots$$

$$= 1$$

$$\left\langle\xi\middle|\textstyle\sum_{ij}\mathbb{P}_{ij}\,\xi\right\rangle = \langle\xi|\mathbb{P}_{12} + \mathbb{P}_{13} + \cdots|\xi\rangle$$

$$= \langle\xi|\mathbb{P}_{12}\xi\rangle + \langle\xi|\mathbb{P}_{13}\xi\rangle + \cdots$$

$$= \int \chi_1^*(\boldsymbol{x}_1)\chi_2^*(\boldsymbol{x}_2)\cdots\chi_N^*(\boldsymbol{x}_N)\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\,d\boldsymbol{x}_1 d\boldsymbol{x}_2\cdots d\boldsymbol{x}_N$$

$$= \int \chi_1^*(\boldsymbol{x}_1)\chi_1(\boldsymbol{x}_2)d\boldsymbol{x}_1 \int \chi_2^*(\boldsymbol{x}_2)\chi_2(\boldsymbol{x}_1)d\boldsymbol{x}_2\cdots = 0$$

Los demás términos también son cero dado que $\delta_{ij} = 0$. Por lo tanto

$$\langle\psi|\psi\rangle = 1$$

$$E = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} = \langle\psi|H|\psi\rangle$$

$$= \langle \mathcal{O}\xi | H | \mathcal{O}\xi \rangle$$
$$= \langle \xi | \mathcal{O}^{\dagger}(H\mathcal{O}\xi) \rangle$$
$$= \langle \xi | \mathcal{O}(H\mathcal{O}\xi) \rangle$$
$$= \langle \xi | H\mathcal{O}\mathcal{O}\xi \rangle$$
$$= \langle \xi | H\mathcal{O}^2\xi \rangle$$
$$= \left\langle \xi \middle| H(N!)^{\frac{1}{2}}\mathcal{O}\xi \right\rangle$$
$$= (N!)^{\frac{1}{2}} \langle \xi | H | \mathcal{O}\xi \rangle$$
$$= (N!)^{\frac{1}{2}} \left\langle \xi \middle| H \middle| (N!)^{-\frac{1}{2}} \left(1 - \sum_{ij} \mathbb{P}_{ij} + \sum_{ijk} \mathbb{P}_{ijk} - \cdots\right)\xi \right\rangle$$
$$= \left\langle \xi \middle| H \middle| \left(1 - \sum_{ij} \mathbb{P}_{ij} + \sum_{ijk} \mathbb{P}_{ijk} - \cdots\right)\xi \right\rangle$$
$$= \langle \xi | H | \xi \rangle - \left\langle \xi \middle| H \middle| \sum_{ij} \mathbb{P}_{ij}\, \xi \right\rangle + \left\langle \xi \middle| H \middle| \sum_{ijk} \mathbb{P}_{ijk}\, \xi \right\rangle - \cdots$$

$$\langle \xi | H | \xi \rangle = \int \chi_1^*(\boldsymbol{x}_1)\chi_2^*(\boldsymbol{x}_2)\cdots\chi_N^*(\boldsymbol{x}_N) \left\{ -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 - \sum_{i=1}^{N}\sum_{\mu=1}^{S} \frac{Z}{|\boldsymbol{r}_i - R_\mu|} + \sum_{i>j}^{N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \right\} \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\, d\boldsymbol{x}_1 d\boldsymbol{x}_2 \cdots d\boldsymbol{x}_N$$

Como $-\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{\mu=1}^{S}\frac{Z}{|\boldsymbol{r}_i - \boldsymbol{r}_\mu|} = \sum_{i=1}^{N}\left\{-\frac{1}{2}\nabla_i^2 - \sum_{\mu=1}^{S}\frac{Z}{|\boldsymbol{r}_i - R_\mu|}\right\}$, entonces el hamiltoniano monoelectrónico está dado por

$$h_i = -\frac{1}{2}\nabla_i^2 - \sum_{\mu=1}^{S} \frac{Z}{|\boldsymbol{r}_i - R_\mu|}$$

$$\therefore \qquad H = h_i + \sum_{i>j}^{N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}$$

Nótese que *i* toma valor único, en consecuencia

$$h_1 = -\frac{1}{2}\nabla_1^2 - \sum_{\mu=1}^{S} \frac{Z}{|\boldsymbol{r}_1 - R_\mu|}$$

$$\langle \xi | H | \xi \rangle = \int \chi_1^*(\boldsymbol{x}_1)\chi_2^*(\boldsymbol{x}_2)\cdots\chi_N^*(\boldsymbol{x}_N) \left\{ \sum_{i=1}^{N} h_i + \sum_{i>j}^{N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \right\} \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2) \cdots \chi_N(\boldsymbol{x}_N) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \cdots d\boldsymbol{x}_N$$

$$= \sum_{i=1}^{N} \int \chi_1^*(\boldsymbol{x}_1) h_1 \chi_1(\boldsymbol{x}_1) d\boldsymbol{x}_1 \int \chi_2^*(\boldsymbol{x}_2)\chi_2(\boldsymbol{x}_2) d\boldsymbol{x}_2 \cdots \int \chi_N^*(\boldsymbol{x}_N)\chi_N(\boldsymbol{x}_N) d\boldsymbol{x}_N$$

$$+ \int \chi_1^*(\boldsymbol{x}_1)\chi_2^*(\boldsymbol{x}_2)\cdots\chi_N^*(\boldsymbol{x}_N)\left\{\frac{1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} + \frac{1}{|\boldsymbol{r}_3 - \boldsymbol{r}_1|} + \cdots + \frac{1}{|\boldsymbol{r}_3 - \boldsymbol{r}_2|} + \frac{1}{|\boldsymbol{r}_4 - \boldsymbol{r}_2|} + \cdots\right\}\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \cdots d\boldsymbol{x}_N$$

$$= \sum_{i=1}^{N} \langle \chi_i(1)|h_1|\chi_i(1)\rangle$$

$$+ \int \chi_1^*(\boldsymbol{x}_1)\chi_2^*(\boldsymbol{x}_2)\left(\frac{1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}\right)\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \int \chi_3^*(\boldsymbol{x}_3)\chi_3(\boldsymbol{x}_3) d\boldsymbol{x}_3 \cdots \int \chi_N^*(\boldsymbol{x}_N)\chi_N(\boldsymbol{x}_N) d\boldsymbol{x}_N$$

$$= \sum_{i=1}^{N} \langle \chi_i(1)|h_1|\chi_i(1)\rangle + \int \chi_1^*(\boldsymbol{x}_1)\chi_2^*(\boldsymbol{x}_2)\left(\frac{1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}\right)\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2) d\boldsymbol{x}_1 d\boldsymbol{x}_2$$

$$= \sum_{i=1}^{N} \langle \chi_i(1)|h_1|\chi_i(1)\rangle + \left\langle \chi_1(1)\chi_2(2)\middle|\frac{1}{r_{12}}\middle|\chi_1(1)\chi_2(2)\right\rangle$$

De tal modo, las integrales se reducen a integrales de dos centros, esto es

$$\langle \xi|H|\xi\rangle = \sum_{i=1}^{N} \langle \chi_i(1)|h_1|\chi_i(1)\rangle + \sum_{i>j} \left\langle \chi_i(1)\chi_j(2)\middle|\frac{1}{r_{12}}\middle|\chi_i(1)\chi_j(2)\right\rangle$$

Obsérvese que en unidades atómicas

$$h_i = -\frac{1}{2}\nabla_i^2 - \sum_{\mu=1}^{S} \frac{1}{r_{i\mu}}$$

$$H = \sum_{i=1}^{N} h_i + \sum_{i>j}^{N} \frac{1}{r_{ij}}$$

y, por lo tanto:

$$\langle \xi|H|\xi\rangle = \left\langle \xi\middle| \textstyle\sum_{i=1}^{N} h_i + \sum_{i>j}^{N} \frac{1}{r_{ij}}\middle|\xi\right\rangle$$

$$= \langle \xi| \textstyle\sum_{i=1}^{N} h_i |\xi\rangle + \left\langle \xi\middle| \sum_{i>j}^{N} \frac{1}{r_{ij}}\middle|\xi\right\rangle$$

$$= \langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)| \textstyle\sum_{i=1}^{N} h_i |\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\rangle$$

$$+ \left\langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\middle| \sum_{i>j}^{N} \frac{1}{r_{ij}}\middle|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\right\rangle$$

$$= \langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)|h_1+h_2+\cdots+h_N|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\rangle$$

$$+\langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\left|\frac{1}{r_{12}}+\frac{1}{r_{13}}+\cdots+\frac{1}{r_{1N}}+\frac{1}{r_{23}}+\frac{1}{r_{24}}+\cdots+\frac{1}{r_{2N}}\right.$$

$$\left.+\frac{1}{r_{34}}+\frac{1}{r_{35}}+\cdots+\frac{1}{r_{3N}}\right|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\rangle$$

$$= \langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)|h_1|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\rangle+\cdots$$

$$+\left\langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\left|\frac{1}{r_{12}}+\frac{1}{r_{13}}+\cdots+\frac{1}{r_{1N}}\right|\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\right\rangle$$

El primer término de esta última igualdad es una integral de un centro ya que solo depende de una coordenada, todas las demás son uno. En el segundo término de esta igualdad, son integrales de dos centros ya que dependen de dos coordenadas, todas las demás no contribuyen.

Ahora para el término de las permutaciones por pares, se tiene que

$$\langle \xi|H|\textstyle\sum_{ij}\mathbb{P}_{ij}\,\xi\rangle = \left\langle \xi\left|\sum_{i=1}^{N}h_i+\sum_{i>j}^{N}\frac{1}{r_{ij}}\right|(\mathbb{P}_{12}+\mathbb{P}_{13}+\cdots+\mathbb{P}_{1N}+\mathbb{P}_{23}+\cdots+)\xi\right\rangle$$

$$= \left\langle \xi\left|\sum_{i=1}^{N}h_i\right|(\mathbb{P}_{12}+\mathbb{P}_{13}+\cdots+\mathbb{P}_{1N}+\mathbb{P}_{23}+\cdots+)\xi\right\rangle$$

$$+\left\langle \xi\left|\frac{1}{r_{12}}+\frac{1}{r_{13}}+\cdots\right|(\mathbb{P}_{12}+\mathbb{P}_{13}+\cdots+\mathbb{P}_{1N}+\mathbb{P}_{23}+\cdots+)\xi\right\rangle$$

$$= \langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)|h_1|\mathbb{P}_{12}\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\rangle$$

$$+\left\langle \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\left|\frac{1}{r_{12}}+\frac{1}{r_{13}}+\cdots\right|\mathbb{P}_{12}\chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\cdots\chi_N(\boldsymbol{x}_N)\right\rangle$$

$$= \int \chi_1^*(1)h_1\chi_2(1)d\boldsymbol{x}_1\int \chi_2^*(2)\chi_1(2)d\boldsymbol{x}_2$$

$$+\int \chi_1^*(1)\chi_2^*(2)\left(\frac{1}{r_{12}}\right)\chi_1(2)\chi_2(1)d\boldsymbol{x}_1d\boldsymbol{x}_2\int \chi_3^*(3)\chi_3(3)d\boldsymbol{x}_3\cdots$$

$$+\int\cdots\left(\frac{1}{r_{12}}\right)\mathbb{P}_{13}\cdots+\int\cdots\left(\frac{1}{r_{12}}\right)\mathbb{P}_{14}\cdots$$

$$+\int \chi_1^*(1)\chi_3^*(3)\left(\frac{1}{r_{13}}\right)\chi_1(3)\chi_3(1)d\boldsymbol{x}_1d\boldsymbol{x}_3\delta_{44}\cdots$$

$$= \int \chi_1^*(1) h_1 \chi_2(1) d\boldsymbol{x}_1 \delta_{12} + \int \chi_1^*(1)\chi_2^*(2)\left(\frac{1}{r_{12}}\right)\chi_1(2)\chi_2(1) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \delta_{33} \cdots$$

$$+0 + 0 + \int \chi_1^*(1)\chi_3^*(3)\left(\frac{1}{r_{13}}\right)\chi_1(3)3(1) d\boldsymbol{x}_1 d\boldsymbol{x}_3 \delta_{44} \cdots$$

$$= 0 + \int \chi_1^*(1)\chi_2^*(2)\left(\frac{1}{r_{12}}\right)\chi_1(2)\chi_2(1) d\boldsymbol{x}_1 d\boldsymbol{x}_2$$

$$+ \int \chi_1^*(1)\chi_3^*(3)\left(\frac{1}{r_{13}}\right)\chi_1(3)3(1) d\boldsymbol{x}_1 d\boldsymbol{x}_3 \cdots + \cdots$$

Por lo tanto, colectando los términos de la energía

$$E = \langle \xi | H | \xi \rangle - \langle \xi | H | \sum_{ij} \mathbb{P}_{ij}\, \xi \rangle + \langle \xi | H | \sum_{ijk} \mathbb{P}_{ijk}\, \xi \rangle - \cdots$$

Se tiene que

$$E = \sum_i \langle \chi_i(1) | h_1 | \chi_i(1) \rangle + \sum_{i>j} \left\langle \chi_i(1)\chi_j(2) \middle| \frac{1}{r_{12}} \middle| \chi_i(1)\chi_j(2) \right\rangle$$

$$- \sum_{i>j} \left\langle \chi_i(1)\chi_j(2) \middle| \frac{1}{r_{12}} \middle| \chi_j(1)\chi_i(2) \right\rangle + \langle \xi | H | \sum_{ijk} \mathbb{P}_{ijk}\, \xi \rangle - \cdots$$

No es complicado mostrar que el último término se anula, de modo que

$$E = \sum_i \langle \chi_i(1) | h_1 | \chi_i(1) \rangle$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\left\langle \chi_i(1)\chi_j(2) \middle| \frac{1}{r_{12}} \middle| \chi_i(1)\chi_j(2) \right\rangle - \left\langle \chi_i(1)\chi_j(2) \middle| \frac{1}{r_{12}} \middle| \chi_j(1)\chi_i(2) \right\rangle \right]$$

Casi todas las moléculas tienen un número par de electrones (2N) y en consecuencia son estados base de capa cerrada. Dado que los orbitales en realidad son espín-orbitales, entonces

$$\left.\begin{matrix} \chi_1(1) \longrightarrow \phi_1(1)\alpha \\ \chi_2(2) \longrightarrow \phi_1(2)\beta \end{matrix}\right\} \quad \left.\begin{matrix} \langle \phi_1(1)\alpha | h_1 | \phi_1(1)\alpha \rangle \\ \langle \phi_1(1)\beta | h_1 | \phi_1(1)\beta \rangle \end{matrix}\right\} \rightarrow \quad 2\sum_i \langle \phi_1(1) | h_1 | \phi_1(1) \rangle$$

ϕ_1 y ϕ_2 son ortogonales. Además, entre las posibilidades de combinar los espinorbitales ponemos dos casos

$$\left\langle \chi_i(1)\chi_j(2)\left|\frac{1}{r_{12}}\right|\chi_j(1)\chi_i(2)\right\rangle = \begin{cases} \left\langle \phi_i(1)\alpha(1)\phi_j(2)\alpha(2)\left|\frac{1}{r_{12}}\right|\phi_j(1)\alpha(1)\phi_i(2)\alpha(2)\right\rangle \\ \text{ó } \left\langle \phi_i(1)\alpha(1)\phi_j(2)\beta(2)\left|\frac{1}{r_{12}}\right|\phi_j(1)\beta(1)\phi_i(2)\alpha(2)\right\rangle \end{cases}$$

$$= \begin{cases} \left\langle \phi_i(1)\phi_j(2)\left|\frac{1}{r_{12}}\right|\phi_j(1)\phi_i(2)\right\rangle \int \alpha^*(1)\alpha(1)\alpha^*(2)\alpha(2)d1d2 \\ \text{ó } \left\langle \phi_i(1)\phi_j(2)\left|\frac{1}{r_{12}}\right|\phi_j(1)\phi_i(2)\right\rangle \int \alpha^*(1)\beta(1)\beta^*(2)\alpha(2)d1d2 \end{cases}$$

$$= \begin{cases} \left\langle \phi_i(1)\phi_j(2)\left|\frac{1}{r_{12}}\right|\phi_j(1)\phi_i(2)\right\rangle \\ 0 \end{cases}$$

La energía de un electrón e^- en presencia de N-1 electrones es

$$E = 2\sum_i \langle \phi_i(1)|h_1|\phi_i(1)\rangle + \sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\left\langle \phi_i(1)\phi_j(2)\left|\frac{1}{r_{12}}\right|\phi_i(1)\phi_j(2)\right\rangle - \left\langle \phi_i(1)\phi_j(2)\left|\frac{1}{r_{12}}\right|\phi_j(1)\phi_i(2)\right\rangle\right]$$

El último término ya se puede identificar como un término de intercambio. Para minimizar la energía E se debe hacer la variación $\delta E = 0$, con la condición de que las funciones $\{\phi_1, \phi_2, \phi_3, \dots, \phi_N\}$ sean ortonormales $\{\langle \phi_i|\phi_j\rangle = \delta_{ij}\}$. El problema entonces es encontrar el conjunto de funciones $\{\phi_1, \phi_2, \phi_3, \dots, \phi_N\}$. Hagamos un paréntesis antes de analizar la variación de la energía *E*, observando la variación en $\langle \phi_i|\phi_j\rangle = \delta_{ij}$, esto es

$$\delta\langle \phi_i|\phi_j\rangle = \delta(\delta_{ij})$$

$$\langle \delta\phi_i|\phi_j\rangle + \langle \phi_i|\delta\phi_j\rangle = 0$$

multiplicando esta expresión por $-2\xi_{ji}$ se tiene que

$$-2\xi_{ji}\langle \delta\phi_i|\phi_j\rangle - 2\xi_{ji}\langle \phi_i|\delta\phi_j\rangle = 0$$

sumando para todas las *i* y *j* se tiene que

$$-2\sum_{ij}\xi_{ji}\langle \delta\phi_i|\phi_j\rangle - 2\sum_{ij}\xi_{ji}\langle \phi_i|\delta\phi_j\rangle = 0$$

Renombrando subíndices esta expresión se puede expresar como

$$-\sum_{ij}\xi_{ji}\langle \delta\phi_i|\phi_j\rangle - \sum_{ij}\xi_{ij}\langle \phi_j|\delta\phi_i\rangle = 0$$

$$\delta E - 2\sum_{ij}\xi_{ji}\langle \delta\phi_i|\phi_j\rangle - 2\sum_{ij}\xi_{ij}\langle \phi_j|\delta\phi_i\rangle = 0$$

Variación de E

$$\delta E = 2\delta\left(\sum_i \langle\phi_i(1)|h_1|\phi_i(1)\rangle\right) + \delta\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\left\langle\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_i(1)\phi_j(2)\right\rangle - \left\langle\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_j(1)\phi_i(2)\right\rangle\right]\right)$$

$$\delta E = 2\sum_i \langle\delta\phi_i(1)|h_1|\phi_i(1)\rangle + 2\sum_i \langle\phi_i(1)|h_1|\delta\phi_i(1)\rangle + \sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\left\langle\delta\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_i(1)\phi_j(2)\right\rangle + 2\left\langle\phi_i(1)\delta\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_i(1)\phi_j(2)\right\rangle + 2\left\langle\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\delta\phi_i(1)\phi_j(2)\right\rangle + 2\left\langle\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_i(1)\delta\phi_j(2)\right\rangle - \left\langle\delta\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_j(1)\phi_i(2)\right\rangle - \left\langle\phi_i(1)\delta\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_j(1)\phi_i(2)\right\rangle - \left\langle\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\delta\phi_j(1)\phi_i(2)\right\rangle - \left\langle\phi_i(1)\phi_j(2)\right|\frac{1}{r_{12}}\left|\phi_j(1)\delta\phi_i(2)\right\rangle\right]$$

Ahora, definiendo a J_i como operador de Coulomb, y a K_i como operador de Intercambio (*i* por *j*).

$$J_i(1)\phi_j(1) = \langle\phi_i(2)|h_1|\phi_i(2)\rangle\phi_j(1)$$

$$K_i(1)\phi_j(1) = \langle\phi_i(2)|h_1|\phi_j(2)\rangle\phi_i(1)$$

Para el operador de Coulomb, lo que siente el electrón e^- en la coordenada de posición (1) del potencial *j* se debe a la distribución de carga del electrón e^- en la coordenada de posición (2) del potencial *i* y la distribución del electrón e^- en la coordenada de posición (2) del potencial *i*. En notación de integrales se tiene

$$J_i(1)\phi_j(1) = \int \frac{\phi_i^*(2)\phi_i(2)}{r_{12}}d\boldsymbol{r}_2$$

$$= \int \frac{(-e)\phi_i^*(2)(-e)\phi_i(2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}d\boldsymbol{r}_2 = V(\boldsymbol{r}_1)$$

Nótese que la interacción entre estos dos electrones solo se da en un punto. Ahora bien, tomando $e = +1$ para el electrón, y en términos de la densidad de potencial

$$\int \frac{(-e)\phi_i^*(2)(-e)\phi_i(2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} d\boldsymbol{r}_2 = \int \frac{\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} d\boldsymbol{r}_2 = V(\boldsymbol{r}_1)$$

donde $V(\boldsymbol{r}_1)$ es el potencial eléctrico local del electrón *i* debido a un e^- descrito por la función ϕ_i ubicada en $\boldsymbol{r}_2$. De la definición de operador de Coulomb, la interacción sobre el electrón e^- respecto al potencial *j* en el punto (1) se debe a la distribución electrónica e^- respecto al potencial *i* en el punto (2).

Mientras que, para el operador de intercambio, lo que siente el electrón e^- en la coordenada de posición (1) del potencial *j* se debe a la distribución de carga del electrón e^- en la coordenada de posición (2) del potencial *i* y la distribución del electrón e^- en la coordenada de posición (2) del potencial *j*. En la notación de integrales se tiene

$$K_i(1)\phi_j(1) = \int \frac{\phi_i^*(2)\phi_j(2)}{r_{12}} d\boldsymbol{r}_2$$

Para K_i se trata de un potencial eléctrico no local que se corresponde a la interacción de *i* y *j* en todo el espacio. De la definición de operador de intercambio, lo que siente el electrón e^- respecto al potencial *j* en el punto (1) se debe a la distribución electrónica e^- respecto al potencial *i* en el punto (2) y a la distribución electrónica del e^- respecto al potencial *j* en todo el espacio.

Reescribiendo los operadores J_i y K_i tenemos lo siguiente

$$J_i(1)\delta\phi_j(1) = \langle\phi_i(2)|h_1|\phi_i(2)\rangle\delta\phi_j(1)$$

$$K_i(1)\delta\phi_j(1) = \langle\phi_i(2)|h_1|\delta\phi_j(2)\rangle\phi_i(1)$$

$$\delta E = 2\sum_{i=1}^{N}\langle\delta\phi_i(1)|h_1|\phi_i(1)\rangle + 2\sum_{i=1}^{N}\langle\phi_i(1)|h_1|\delta\phi_i(1)\rangle$$
$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\langle\delta\phi_i(1)|J_j(1)|\phi_i(1)\rangle - \langle\delta\phi_i(1)|K_j(1)|\phi_i(1)\rangle\right]$$
$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\langle\delta\phi_j(2)|J_i(2)|\phi_j(2)\rangle - \langle\delta\phi_j(2)|K_i(2)|\phi_j(2)\rangle\right]$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\langle\phi_i(1)|J_j(1)|\delta\phi_i(1)\rangle-\langle\phi_j(2)|K_i(2)|\delta\phi_j(2)\rangle\right]$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\langle\phi_j(2)|J_i(2)|\delta\phi_j(2)\rangle-\langle\phi_i(1)|K_j(1)|\delta\phi_i(1)\rangle\right]$$

Al hacer que la coordenada (2) sea igual a la (1), y cambiar *i* por *j* se tiene lo siguiente:

$$\delta E=2\left\{\sum_{i=1}^{N}\langle\delta\phi_i(1)|h(1)|\phi_i(1)\rangle+\sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\langle\delta\phi_i(1)|J_j(1)|\phi_i(1)\rangle-\langle\delta\phi_i(1)|K_j(1)|\phi_i(1)\rangle\right]\right\}$$

$$+2\left\{\sum_{i=1}^{N}\langle\phi_i(1)|h(1)|\delta\phi_i\rangle+\sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\langle\phi_i(1)|J_j(1)|\delta\phi_i(1)\rangle-\langle\phi_i(1)|K_j(1)|\delta\phi_i(1)\rangle\right]\right\}$$

Dado que $h(1)$, J_j, y K_j son operadores hermitianos, esto es

$$\langle\phi_i(1)|h(1)|\delta\phi_i(1)\rangle=\langle h(1)\phi_i(1)|\delta\phi_i(1)\rangle$$

$$=\langle\delta\phi_i(1)|h(1)|\phi_i(1)\rangle^*$$

$$\langle\phi_i(1)|J_j(1)|\delta\phi_i(1)\rangle=\langle J_j(1)\phi_i(1)|\delta\phi_i(1)\rangle$$

$$=\langle\delta\phi_i(1)|J_j(1)|\phi_i(1)\rangle^*$$

$$\langle\phi_i(1)|K_j(1)|\delta\phi_i(1)\rangle=\langle K_j(1)\phi_i(1)|\delta\phi_i(1)\rangle$$

$$=\langle\delta\phi_i(1)|K_j(1)|\phi_i(1)\rangle^*$$

Entonces

$$\delta E-2\sum_{ij}\xi_{ji}\langle\delta\phi_i|\phi_j\rangle-2\sum_{ij}\xi_{ij}\langle\phi_i|\delta\phi_j\rangle^*=0$$

Nótese que $\langle \delta\phi_i | \phi_j \rangle = \langle \phi_i | \delta\phi_j \rangle^*$ por lo tanto

$$2\sum_{i=1}^{N}\left\{\left\langle \delta\phi_i(1)\middle| h(1) + \sum_{j=1}^{N}\left(2J_j(1) - K_j(1)\right)\middle|\phi_i(1)\right\rangle - \sum_{i=1}^{N}\xi_{ji}\langle \delta\phi_i|\phi_j\rangle\right\}$$

$$+2\sum_{i=1}^{N}\left\{\left\langle \delta\phi_i(1)\middle| h(1) + \sum_{j=1}^{N}\left(2J_j(1) - K_j(1)\right)\middle|\phi_i(1)\right\rangle^* - \sum_{i=1}^{N}\xi_{ji}\langle \delta\phi_i|\phi_j\rangle^*\right\} = 0$$

Se cumple para toda $\delta\phi_i$. Estas dos ecuaciones se pueden escribir como

$$2\sum_{i=1}^{N}\left\{\int\left[h(1) + \sum_{j=1}^{N}\left(2J_j(1) - K_j(1)\right)\right]\phi_i(1) - \sum_{j=1}^{N}\int \xi_{ji}\,\phi_j\right\}\delta\phi_i^*(1)d\boldsymbol{r}_1 = 0$$

$$\sum_{i=1}^{N}\left\{\int\left[h^*(1) + \sum_{j=1}^{N}\left(2J_j^*(1) - K_j^*(1)\right)\right]\phi_i^*(1) - \sum_{j=1}^{N}\int \xi_{ji}\,\phi_j^*\right\}\delta\phi_i(1)d\boldsymbol{r}_1 = 0$$

por lo tanto, lo que debe ser cero es lo que se encuentra entre paréntesis.

$$\left[h(1) + \sum_{j=1}^{N}\left(2J_j(1) - K_j(1)\right)\right]\phi_i - \sum_{i=1}^{N}\xi_{ji}\phi_j = 0$$

$$\left[h^*(1) + \sum_{j=1}^{N}\left(2J_j^*(1) - K_j^*(1)\right)\right]\phi_i^* - \sum_{i=1}^{N}\xi_{ji}\phi_j^* = 0$$

Ahora, tomando el complejo conjugado de esta última ecuación y multiplicarla por menos se tiene que

$$-\left(h(1) + \sum_{j=1}^{N}\left(2J_j(1) - K_j(1)\right)\right)\phi_i + \sum_{i=1}^{N}\xi_{ij}^*\phi_j = 0.$$

Ahora tenemos dos ecuaciones que al sumarlas se obtiene

$$-\sum_{i=1}^{N}\xi_{ji}\phi_j + \sum_{i=1}^{N}\xi_{ij}^*\phi_j = 0$$

de donde $$\sum_{i=1}^{N}\left(\xi_{ij}^{*}-\xi_{ji}\right)\phi_j = 0$$

$$\therefore \qquad \xi_{ij}^{*} = \xi_{ji}$$

Esto es, ξ_{ji} define a una matriz hermitiana

$$\left[h(1)+\sum_{j=1}^{N}\Big(2J_j(1)-K_j(1)\Big)\right]\phi_i-\sum_{i=1}^{N}\xi_{ji}\phi_i=0 \qquad \forall i=1,2,\dots,N$$

Desarrollando, se tiene el sistema

$$\left[h(1)+\sum_{j=1}^{N}\Big(2J_j(1)-K_j(1)\Big)\right]\phi_1-\sum_{i=1}^{N}\xi_{ji}\phi_1=0$$

$$\vdots \qquad\qquad \vdots \qquad \vdots$$

$$\left[h(1)+\sum_{j=1}^{N}\Big(2J_j(1)-K_j(1)\Big)\right]\phi_N-\sum_{i=1}^{N}\xi_{ji}\phi_N=0$$

A este sistema de ecuaciones integrodiferenciales acopladas se les conoce como ecuaciones de Hartree-Fock (*H-F*). Si se definen las matrices

$$\mathbb{N}=\begin{pmatrix}\phi_1\\ \phi_2\\ \vdots\\ \phi_N\end{pmatrix} \qquad y \qquad \mathbb{E}=\xi_{ji}$$

y al tomar

$$\mathbb{F}=h+\sum_{j=1}^{N}\left(2J_j-K_j\right)$$

y notar que el producto $\xi_{ji}\phi_i$ tiene índices repetidos lo cual indica sumatoria, la ecuación *H-F* se puede escribir como

$$\mathbb{F}\mathbb{N}=\mathbb{N}\mathbb{E}$$

Resumiendo, la función de onda antisimétrica más simple que puede ser usada para describir el estado base de un sistema de N-electrones es el determinante de Slater, que en notación de Dirac es

$$|\Psi_0\rangle=|\chi_1\chi_2\cdots\chi_N\rangle$$

El principio variacional propone que la mejor función de onda de esta forma funcional es la única que da la energía más baja posible

$$E_0=\langle\Psi_0|H_{electrónico}|\Psi_0\rangle$$

La flexibilidad variacional en la función de onda está en la elección de los espín-orbitales. La ecuación de Hartree-Fock se deriva minimizando E_0 y buscando determinar los espín-orbitales óptimos. La ecuación monoelectrónica de Hartree-Fock es una ecuación de eigenvalores de la forma

$$f(i)\chi(\boldsymbol{x}_i) = \varepsilon\chi(\boldsymbol{x}_i) \tag{78}$$

donde ε es la energía del espín-orbital $\chi(\boldsymbol{x}_i)$, además de ser el parámetro variacional por minimizar usando la relación de Rayleigh: $\varepsilon = \frac{\int \psi^*(\boldsymbol{x},\boldsymbol{R})H\psi(\boldsymbol{x},\boldsymbol{R})d\boldsymbol{x}}{\int \psi^*(\boldsymbol{x},\boldsymbol{R})\psi(\boldsymbol{x},\boldsymbol{R})d\boldsymbol{x}}$, $f(i)$ es un operador efectivo de un electrón conocido como operador de *Fock*, de la forma

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_k \frac{Z_k}{r_{ik}} + v^{HF}(i)$$

donde $v^{HF}(i)$ es el potencial promedio experimentado por el *i*-ésimo electrón debido a la presencia de los demás electrones. Las N ecuaciones monoelectrónicas para los orbitales $\chi_1(1), \chi_2(2), \cdots, \chi_N(N)$ son un conjunto de ecuaciones acopladas integrodiferenciales no lineales. El potencial de Hartree-Fock $v^{HF}(i)$, o equivalentemente el campo visto por el *i*-ésimo electrón, depende de los espines orbitales de los demás electrones (*i.e.* el operador de *Fock* depende de sus eigenfunciones) y está dado por

$$v^{HF}(i) = \sum_b J_b(i) - K_b(i)$$

donde el operador de Coulomb J_b, y el operador de intercambio K_b son:

$$J_b(i)\chi_a(i) = \left\{\int d\boldsymbol{x_j}\, \chi_b^*(j) r_{ij}^{-1} \chi_b(j)\right\}\chi_a(i)$$

$$K_b(i)\chi_a(i) = \left\{\int d\boldsymbol{x_j}\, \chi_b^*(j) r_{ij}^{-1} \chi_a(j)\right\}\chi_b(i)$$

Los operadores de Coulomb y de intercambio están en términos de los espín-orbitales, el operador de Coulomb toma la repulsión Coulombiana entre electrones y el operador de intercambio representa la modificación de esta energía que puede ser referida a los efectos de correlación de espín en coordenadas espaciales.

La ecuación de Hartree-Fock para el espín-orbital $\chi_a(1)$ viene dada por

$$f_1\chi_a(1) = \varepsilon_a\chi_a(1), \tag{79}$$

donde ε_a es la energía orbital del espín-orbital $\chi_a(1)$, y f_1 el operador de Fock es

$$f_1 = h_1 + \sum_b \{J_b(1) - K_b(1)\}$$

donde

$$h_1 = -\frac{1}{2}\nabla_1^2 - \sum_{k=1}^{S} \frac{Z_k}{r_{ik}}$$

Nótese que debido a que $J_a(1)\chi_a(1) = K_a(1)\chi_a(1)$ la sumatoria anterior incluye las contribuciones de todos los espín-orbitales χ_b excepto el espín-orbital χ_a que está siendo calculado.

Cada espín-orbital debe obtenerse resolviendo una ecuación de la forma de la ecuación (79), con el correspondiente operador de Fock *f*(*i*). No obstante, debido a que *f*(*i*) depende de los *N*-1 electrones, y dado que el sistema es no lineal, para conjuntar las ecuaciones de HF se deben conocer de antemano las soluciones. Se adopta el esquema iterativo SCF del que habíamos hablado previamente, deteniendo este proceso cuando las soluciones son autoconsistentes. De aquí el nombre método Hartree-Fock de campo autoconsistente o HF-SCF, por sus siglas en inglés. Se formula un conjunto de espín-orbitales de prueba y se usa para formar el operador de Fock, entonces las ecuaciones de HF se resuelven para obtener un nuevo conjunto de espín-orbitales que se usan para construir un operador corregido de Fock. El ciclo de cálculo y reformulación es repetido hasta que sea satisfecho un cierto criterio de convergencia. El criterio usado en este caso es que el determinante de la ecuación secular de las ecuaciones de Roothaan sea igual a cero.

La idea básica del método SCF (*self consistent field*) es simple, haciendo una estimación inicial en los espín-orbitales se puede calcular el campo promedio (*i.e.*, $v^{HF}(i)$) visto por cada electrón, y entonces resolver la ecuación de eigenvalores (78) para un nuevo conjunto de espín-orbitales. Al usar estos nuevos espín-orbitales se pueden obtener nuevos campos y repetir el procedimiento hasta que se alcance la autoconsistencia (*i.e.*, hasta que los campos ya no cambien más y los espín-orbitales usados para construir el operador de *Fock* sean los mismos que sus eigenfunciones).

La solución al problema de eigenvalores (78) de Hartree-Fock da un conjunto $\{\chi_k\}$ de espín-orbitales ortonormales de Hartree-Fock con energías orbitales $\{\varepsilon_k\}$. Los N espín-orbitales con las energías más bajas son llamados los espín-orbitales *ocupados* u *hoyos*. El

determinante de Slater formado por estos orbitales es el estado base de Hartree-Fock de la función de onda y es la mejor aproximación variacional al estado base del sistema reducido a un solo determinante. Etiquetemos a los espín-orbitales ocupados por los índices $a, b, c, \cdots (i.e., \chi_a, \chi_b, \chi_c, \cdots)$. Los miembros restantes del conjunto $\{\chi_k\}$ se llaman espín-orbitales *virtuales, desocupados*, o *partículas*, los cuales se etiquetarán por los índices $r, s, t, \cdots (i.e., \chi_r, \chi_s, \chi_t, \cdots)$.

En principio, existe un número infinito de soluciones de la ecuación (78) de Hartree-Fock y un número infinito de espín-orbitales virtuales. En la práctica la ecuación de Hartree Fock se resuelve introduciendo un conjunto finito de funciones base $\{\phi_\mu(\boldsymbol{r}) \mid \mu = 1,2, \ldots, K\}$, cuyas partes espaciales de los espín-orbitales con la función de espín α se pueden expandir en términos del conjunto de funciones $\{\phi_\mu\}$. Las partes espaciales de los espín-orbitales con el espín β se pueden expandir de la misma manera y ambas expansiones se sustituyen en el problema de eigenvalores (78), para obtener ecuaciones de eigenvalores matriciales para los coeficientes de expansión. Estas ecuaciones matriciales son las ecuaciones de Roothaan. Se debe entender que usar un conjunto de bases de K funciones espaciales $\{\phi_\mu\}$ conduce a un conjunto de 2K espín-orbitales (K con espín α y K con espín β). Esto conduce a un conjunto de N espín-orbitales ocupados $\{\chi_a\}$ y un conjunto complementario de 2K-N espín-orbitales virtuales o desocupados $\{\chi_r\}$. Un solo determinante de Slater formado por el conjunto $\{\chi_a\}$ es el estado base variacional de Hartree-Fock, para el cual usaremos el símbolo Ψ_0 o $|\Psi_0\rangle = |\chi_1\chi_2 \cdots \chi_a\chi_b \cdots \chi_N\rangle$ El determinante del estado base de Hartree-Fock, $|\chi_1\chi_2 \cdots \chi_a\chi_b \cdots \chi_N\rangle$ ha sido bien representado esquemáticamente en Szabo y Ostlund (1996). Los 2K espín-orbitales de Hartree-Fock se ordenan de acuerdo con su energía, y se desprecian posibles degeneraciones.

En el conjunto más grande y completo de funciones base $\{\phi_\mu\}$, mientras más grande sea el grado de flexibilidad en la expansión para los espín-orbitales, menor será el valor esperado $E_0 = \langle\Psi_0|H_{electrónico}|\Psi_0\rangle$. Conjuntos de bases más y más grandes mantendrán descendiendo la energía E_0 de Hartree-Fock hasta un límite alcanzado, llamado el límite de Hartree-Fock. En la práctica cualquier valor finito de K conducirá a una energía un poco arriba del límite de Hartree-Fock.

Solución de Roothaan a ecuación H-F

A continuación se demuestra que cualquier conjunto M de funciones ϕ' linealmente independientes se obtiene mediante una transformación lineal unitaria de las funciones φ_i que también satisfacen las ecuaciones *H-F*.

Para llevar a cabo esto, se toma la matriz $\mathbb{N}$ de dimensión $1 \times M$ que satisface la ecuación *H-F*, y una matriz $\mathbb{N}'$ de dimensión $1 \times M$ de orbitales φ'

$$\mathbb{N}' = (\varphi_1', \varphi_2', \dots, \varphi_M')$$

Estas matrices se relacionan mediante la matriz unitaria $\mathbb{U}$

$$\mathbb{N}' = \mathbb{N}\mathbb{U}$$

como $\mathbb{U}$ es unitaria, se tiene que

$$\mathbb{U}^\dagger\mathbb{U} = \mathbf{1} = \mathbb{U}\mathbb{U}^\dagger$$

Es decir

$$\mathbb{N}'\mathbb{U}^\dagger = \mathbb{N}\mathbb{U}\mathbb{U}^\dagger$$

$$\mathbb{N} = \mathbb{N}'\mathbb{U}^\dagger$$

Ahora, sustituyendo en la ecuación *H-F*, obtenemos

$$\mathbb{F}\mathbb{N}'\mathbb{U}^\dagger = \mathbb{N}'\mathbb{U}^\dagger\mathbb{E}$$

$$\mathbb{F}\mathbb{N}'\mathbb{U}^\dagger\mathbb{U} = \mathbb{N}'\mathbb{U}^\dagger\mathbb{E}\mathbb{U}$$

$$\mathbb{F}\mathbb{N}' = \mathbb{N}'\mathbb{U}^\dagger\mathbb{E}\mathbb{U}$$

Como $\mathbb{E}$ es una matriz hermitiana, existe una matriz unitaria $\mathbb{U}$ para la cual $\mathbb{U}^\dagger\mathbb{E}\mathbb{U}$ es una matriz diagonal con elementos ϵ_i reales

$$\mathbb{F}\mathbb{N}' = (\varphi_1', \varphi_2', \dots, \varphi_M')\begin{pmatrix} \epsilon_{11} & 0 & \cdots & 0 \\ 0 & \epsilon_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon_{MM} \end{pmatrix}$$

$$\left[h(1) + \sum_{j=1}^{N}\Big(2J_j(1) - K_j(1)\Big)\right](\varphi_1', \varphi_2', \dots, \varphi_N') = (\varphi_1', \varphi_2', \dots, \varphi_N')\begin{pmatrix} \epsilon_{11} & 0 & \cdots & 0 \\ 0 & \epsilon_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon_{NN} \end{pmatrix}$$

$$\left[h(1) + \sum_{j=1}^{N}\Big(2J_j(1) - K_j(1)\Big)\right]\varphi_1' = \varphi_1'\epsilon_{11}$$

Dado que el operador monoelectrónico de Fock es

$$F = h + \sum_{j=1}^{N} \left(2J_j - K_j\right)$$

Entonces

$$F\varphi_i = \epsilon_i \varphi_i$$

$$F|\varphi_i\rangle = \epsilon_i |\varphi_i\rangle$$

$$\begin{matrix} Sistema\ de \\ ecuaciones \\ diferenciales \end{matrix} \begin{cases} F\varphi_1 = \epsilon_1\varphi_1 \\ F\varphi_2 = \epsilon_2\varphi_2 \\ \vdots \\ F\varphi_N = \epsilon_N\varphi_N \end{cases}$$

$$F(\boldsymbol{r}_1)\varphi_N(\boldsymbol{r}_1) = \epsilon_N\varphi_N(\boldsymbol{r}_1)$$

Estas ecuaciones se pueden escribir como

$$h(1)\varphi_i + \sum_{i=1}^{N} \int \frac{\phi_i^*(2)\phi_i(2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} d\boldsymbol{r}_2 \varphi_i(1) - \sum_{i=1}^{N} \int \frac{\phi_j^*(2)\phi_i(2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} d\boldsymbol{r}_2 \varphi_j(1) = \epsilon_i \varphi_i(1)$$

$$h(1) = -\frac{\nabla_i^2}{2} - \sum_{\mu=1}^{S} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_\mu|}$$

Roothaan propuso que los orbitales de Hartree-Fock se representen como combinaciones lineales de un conjunto completo de funciones conocidas $\varphi_i = \sum_p C_{pi}\phi_p$, es decir, el orbital φ_i en términos de un conjunto adecuado de funciones base $\{\phi_p\}$ para atacar este sistema de ecuaciones integrodiferenciales. Entonces, se propone

$$F|\varphi_i\rangle = \epsilon_i |\varphi_i\rangle$$

Y multiplicando por $\langle\phi_s|$ se tiene

$$\begin{aligned} \langle\phi_s|F|\varphi_i\rangle &= \langle\phi_s|\epsilon_i|\varphi_i\rangle \\ &= \epsilon_i \langle\phi_s| \textstyle\sum_p C_{pi} |\phi_p\rangle \\ &= \epsilon_i \sum_{p=1}^{N} C_{pi} \langle\phi_s|\phi_p\rangle \\ &= \epsilon_i \sum_{p=1}^{N} C_{pi} S_{sp} \\ &= \epsilon_i \sum_{p=1}^{N} S_{sp} C_{pi} \\ &= \epsilon_i (SC)_{si} \end{aligned}$$

Nótese que $S_{sp} = \langle \phi_s | \phi_p \rangle$ es la integral de traslape. Tomemos ahora el camino directo

$$\langle \phi_s | F | \phi_p \rangle = \langle \phi_s | h + \sum_{j=1}^{N} (2J_j - K_j) | \phi_p \rangle$$

$$= \langle \phi_s | h | \phi_p \rangle + \langle \phi_s | \sum_{j=1}^{N} (2J_j - K_j) | \phi_p \rangle$$

Dado que

$$J_j \phi_p = \left(\int \frac{\varphi_j^*(2)\varphi_j(2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} d\boldsymbol{r}_2 \right) \phi_p$$

$$K_j \phi_p = \left(\int \frac{\varphi_j^*(2)\phi_p(2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} d\boldsymbol{r}_2 \right) \varphi_j(1)$$

donde ϕ son bases, y φ son orbitales. Entonces

$$\langle \phi_s | F | \phi_p \rangle = \langle \phi_s | h(1) | \phi_p \rangle + \sum_{j=1}^{N} \left\langle \phi_s(1)\varphi_j(2) \left| \frac{1}{r_{12}} \right| \varphi_j(2)\phi_p(1) \right\rangle$$

$$- \sum_{j=1}^{N} \left\langle \phi_s(1)\varphi_j(2) \left| \frac{1}{r_{12}} \right| \phi_p(2)\varphi_j(1) \right\rangle$$

Dado que $\varphi_j = \sum_\ell C_{\ell j}\phi_\ell$, $|\varphi_j\rangle = \sum_\ell C_{\ell j}|\phi_\ell\rangle$, y $\varphi_j = \sum_t C_{tj}\phi_t$ entonces

$$\langle \phi_s | F | \phi_p \rangle = \langle \phi_s | h(1) | \phi_p \rangle + \sum_{j=1}^{N} \left\langle \phi_s(1) \left(\sum_\ell C_{\ell j} | \phi_\ell \rangle \right) \left| \frac{1}{r_{12}} \right| \left(\sum_t C_{tj}\phi_t \right) \phi_p(1) \right\rangle$$

$$- \sum_{j=1}^{N} \left\langle \phi_s(1) \left(\sum_\ell C_{\ell j} | \phi_\ell \rangle \right) \left| \frac{1}{r_{12}} \right| \phi_p(2) \left(\sum_t C_{tj}\phi_t \right) \right\rangle$$

$$\langle \phi_s | F | \phi_p \rangle = \langle \phi_s | h(1) | \phi_p \rangle + \sum_{j=1}^{N} \sum_\ell \sum_t C_{\ell j}^* C_{tj} \left\langle \phi_s \phi_\ell \left| \frac{1}{r_{12}} \right| \phi_t \phi_p \right\rangle$$

$$- \sum_{j=1}^{N} \sum_\ell \sum_t C_{\ell j}^* C_{tj} \left\langle \phi_s(1)\phi_\ell(2) \left| \frac{1}{r_{12}} \right| \phi_p(2)\phi_t(1) \right\rangle$$

Nótese que se han encontrado integrales de dos centros, como

$$\left\langle \phi_s \phi_\ell \left| \frac{1}{r_{12}} \right| \phi_t \phi_p \right\rangle = \iint \phi_s^*(1)\phi_\ell^*(2) \left[\frac{1}{r_{12}} \right] \phi_t(1)\phi_p(2) d\boldsymbol{r}_1 d\boldsymbol{r}_2$$

Por lo tanto, la matriz

$$\begin{aligned}\langle \phi_s|F|\varphi_i\rangle &= \langle \phi_s|F|\textstyle\sum_p C_{pi}\phi_p\rangle \\ &= \sum_p C_{pi}\,\langle \phi_s|F|\phi_p\rangle \\ &= \sum_p C_{pi}\,(F)_{sp} \\ &= \sum_p (F)_{sp}C_{pi} \\ &= (FC)_{si}\end{aligned}$$

Dado que $F\varphi_i = \epsilon_i\varphi_i$ y por el resultado anterior, entonces

$$(FC)_{si} = \epsilon_i(SC)_{si}$$

donde

$$\epsilon_i = \begin{pmatrix} \epsilon_1 & 0 & \cdots & 0 \\ 0 & \epsilon_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon_n \end{pmatrix}$$

Por lo tanto, en notación matricial

$$\mathbb{FC} = \mathbb{SC}E$$

Roothaan propuso esta representación matricial de las ecuaciones de Hartree-Fock y que los orbitales se representaran como combinaciones lineales de un conjunto completo de funciones definidas como *funciones base* espaciales, con lo cual la ecuación diferencial espacial se puede expresar de la siguiente manera:

$$f(\boldsymbol{r}_1)\psi_i(\boldsymbol{r}_1) = \epsilon_i\psi_i(\boldsymbol{r}_1) \qquad (80)$$

la cual se convierte en un conjunto de ecuaciones algebraicas que se puede resolver con métodos matriciales comunes. Más concisamente, introduciendo un conjunto de K funciones base $\{\phi_\mu(\boldsymbol{r}) \mid \mu = 1,2,\dots,K\}$ conocidas y sustituyendo los orbitales moleculares de la expansión lineal

$$\psi_i = \sum_{\mu=1}^{K} C_{\mu i}\,\phi_\mu \qquad i = 1,2,\dots,K \qquad (81)$$

Al sustituir esta expansión lineal en la ecuación de Hartree-Fock (80) y usar el índice *v* se obtiene

$$f(1)\sum_\nu C_{\nu i}\,\phi_\nu(1) = \varepsilon_i \sum_\nu C_{\nu i}\,\phi_\nu(1)$$

cuando se multiplica por $\phi_\mu^*(1)$ a la izquierda e integra se consigue cambiar la ecuación integro-diferencial en una ecuación matricial

$$\sum_\nu C_{\nu i} \int d\boldsymbol{r}_1\, \phi_\mu^*(1) f(1) \phi_\nu(1) = \varepsilon_i \sum_\nu C_{\nu i} \int d\boldsymbol{r}_1\, \phi_\mu^*(1) \phi_\nu(1)$$

Ahora bien, definiendo la matriz de traslape S y la matriz de Fock **F** como

$$S_{\mu\nu} = \int d\boldsymbol{r}_1\, \phi_\mu^*(1) \phi_\nu(1) \quad \text{y} \quad F_{\mu\nu} = \int d\boldsymbol{r}_1\, \phi_\mu^*(1) f(1) \phi_\nu(1)$$

$$S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle \quad \text{y} \quad F_{\mu\nu} = \langle \phi_\mu | f | \phi_\nu \rangle \tag{82}$$

Respectivamente, las ecuaciones de Hartree-Fock se pueden escribir como

$$\sum_\nu F_{\mu\nu}\, C_{\nu i} = \varepsilon_i \sum_\nu S_{\mu\nu}\, C_{\nu i} \qquad i = 1,2,\ldots,K \tag{83}$$

Estas son las ecuaciones de Roothaan, que se pueden escribir como una simple ecuación de matrices $\mathbf{FC} = \mathbf{SC}\varepsilon$ donde **C** es una matriz cuadrada de dimensión K×K de la expansión de los coeficientes $C_{\nu i}$ y ε es una matriz diagonal de las energías orbitales, con ceros afuera de la diagonal como ya se mostró anteriormente. Las ecuaciones (83) forman un conjunto de K ecuaciones simultáneas lineales homogéneas en las K incógnitas $C_{\nu i}$, ν=1,2, …,K, que describen los MO ψ en (81). Para una solución no trivial debemos tener

$$det\left(F_{\mu\nu} - \varepsilon_i S_{\mu\nu}\right) = 0 \tag{84}$$

Esta es una ecuación secular cuyas raíces dan las energías orbitales ε_i. Las ecuaciones de Roothaan (83) deben resolverse por un proceso iterativo, puesto que las integrales $F_{\mu\nu}$ dependen de los orbitales ψ_i, que a su vez dependen de los coeficientes $C_{\nu i}$ desconocidos.

Uno empieza con expresiones de prueba para los MO ocupados como combinaciones lineales de las funciones base, como en la ecuación (81). Este conjunto inicial de MO's se usa para computar el operador de Fock. Los elementos de matriz (82) se computan y se resuelve la ecuación secular (84) para producir un conjunto inicial de ε_i's. Estas ε_i's se emplean para resolver (83) un conjunto mejorado de coeficientes, dando un conjunto mejorado de MO's, los cuales sirven para computar una f mejorada, y así sucesivamente. Se continúa hasta que no ocurren más mejoras en los coeficientes y energías de los MO de un ciclo al siguiente, o si se valida un criterio de convergencia a una tolerancia apropiada. Este método se conoce como método de campo autoconsistente SCF (por sus siglas en inglés).

Funciones base

Las funciones de Slater tienden a infinito cerca del cero, esto se corrige usando funciones gaussianas. En un conjunto de bases contraído de funciones gaussianas, cada función base es una combinación lineal fija de funciones gaussianas primitivas. Generalmente, cada MO se escribe como una combinación lineal de funciones de un electrón (orbitales) centrados en cada átomo.

Por ejemplo, para moléculas diatómicas una buena opción es usar funciones de Slater del tipo

$$\frac{(2\zeta/a_0)^{n+1/2}}{[(2n)!]^{1/2}} r^{n-1} e^{-\zeta r/a_0} Y_\ell^m(\theta,\varphi)$$

para los orbitales atómicos (AO), que se conocen como STO (*slater type orbitals*). El conjunto de todas estas funciones con n, ℓ, m enteros y con ζ tomando todos los valores positivos posibles forma un conjunto completo. El parámetro ζ se llama *exponente orbital*. Por ejemplo, una función 1s depende de $e^{-\zeta r/a_0}$, una función *2p* de $re^{-\zeta r/a_0}$, una *3d* de $r^2 e^{-\zeta r/a_0}$, y así sucesivamente. Para tener un conjunto completo de funciones base AO, se necesita un número infinito de orbitales de Slater, pero la función de onda molecular verdadera de Hartree-Fock puede ser adecuadamente aproximada con un número razonablemente pequeño de orbitales de Slater cuidadosamente elegidos. Un *conjunto de bases mínimo* para un cálculo molecular SCF consiste en una sola función base para cada capa interna AO y cada capa de valencia de AO de cada átomo. Un conjunto de bases extendido es un conjunto que es más grande que un conjunto mínimo. Cálculos SCF con un conjunto-de-bases-mínimo son más fáciles que cálculos con un conjunto-de-bases extendido, pero estos últimos son considerablemente más exactos.

Solo funciones tipo-Slater o tipo-Gaussiano son de uso común. Si uno usa un conjunto muy pequeño de funciones por átomo, entonces las funciones tipo-Slater dan energías definitivamente superiores. Conforme el número de funciones por átomo incrementa, la clara superioridad de las funciones tipo-Slater disminuye, así como también la habilidad del conjunto de bases para abarcar el espacio de la función. Uno tiene que considerar por razones prácticas el tiempo requerido para evaluar las integrales moleculares. La mayoría de los

cálculos poliatómicos usa orbitales tipo gaussiano (GTO) debido a la velocidad con la cual se pueden evaluar las integrales. Las funciones gaussianas cartesianas son de la forma

$$\theta_{\ell mn}(\mathbf{r}_1 - \mathbf{r}_c) = (x_1 - x_c)^{\ell}(y_1 - y_c)^{m}(z_1 - z_c)^{n} e^{-\alpha|\mathbf{r}_1 - \mathbf{r}_c|^2}$$

Donde (x_c,y_c,z_c) son las coordenadas cartesianas del centro de la gaussiana en $\mathbf{r}_c$; ℓ, m, y n son enteros no-negativos; y α es un exponente positivo. Si $\ell = m = n = 0$ entonces la gaussiana cartesiana es una gaussiana tipo-s; cuando $\ell+m+n=1$ es una tipo-p, cuando $\ell+m+n=2$ se trata de una tipo-d, y así sucesivamente. La ventaja central de los GTO es que el producto de dos gaussianas en diferentes centros es equivalente a una sola función gaussiana centrada en un punto entre los dos centros. Por lo tanto, integrales de dos-electrones en tres y cuatro diferentes centros atómicos pueden reducirse a integrales sobre dos diferentes centros, que son mucho más fáciles de computar. Sin embargo, una desventaja de usar los GTO es, por ejemplo, que un orbital atómico hidrogenoide 1s tiene un pico en el núcleo atómico, un $n=1$ STO también lo tiene allí, pero una GTO no. Dado que una GTO da una representación más pobre de los orbitales en el núcleo atómico se debe usar una base más grande para conseguir una precisión comparable a la obtenida de STO. Para aliviar este problema se usa un conjunto de bases contraído de funciones gaussianas, en el cual uno deja que cada función base sea una combinación lineal fija (contracción) de funciones Gaussianas primitivas, g. Una contracción así tiene la forma

$$\phi_{\mu}^{CGF}(\mathbf{r} - \mathbf{R}_{\boldsymbol{A}}) = \sum_{p=1}^{L} d_{p\mu}\, \mathrm{g}_p\left(\alpha_{p\mu}, \mathbf{r} - \mathbf{R}_{\boldsymbol{P}}\right) \qquad (85)$$

donde $\alpha_{p\mu}$ y $d_{p\mu}$ son los exponentes y coeficientes de contracción y L es la longitud de la contracción. Los orígenes $\mathbf{R}_P$ de las primitivas en (85) son casi siempre iguales a $\mathbf{R}_A$. Las funciones Gaussianas primitivas normalizadas son del tipo 1s, 2p, 3d, . . . ,

$$\mathrm{g}_{1s}(\alpha, \mathbf{r}) = (8\alpha^3/\pi^3)^{1/4} e^{-\alpha r^2}$$

$$\mathrm{g}_{2p}(\alpha, \mathbf{r}) = (128\,\alpha^5/\pi^3)^{1/4} x e^{-\alpha r^2}$$

$$\mathrm{g}_{3d_{xy}}(\alpha, \mathbf{r}) = (2048\,\alpha^7/\pi^3)^{1/4} xy e^{-\alpha r^2}$$

El uso de gaussianas contraídas en lugar de las primitivas reduce el número de coeficientes desconocidos $d_{p\mu}$ a ser determinados en los cálculos HF. Por ejemplo, si cada gaussiana contraída se compone de tres primitivas de un conjunto de 30 funciones base primitivas, entonces mientras que la expansión en la ecuación (81) involucra 30 coeficientes desconocidos $d_{p\mu}$, la correspondiente expansión usando la ecuación (85) tiene solo 10 coeficientes desconocidos. Esta disminución en el número de coeficientes conduce a un ahorro en tiempo de cómputo potencialmente grande con poca pérdida de precisión si las gaussianas contraídas están bien elegidas.

Densidad electrónica

La función antisimétrica más simple para representar al estado fundamental de una molécula o de un sistema molecular con N electrones, es un solo determinante de Slater construido con N espín-orbitales

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_N) = (N!)^{-\frac{1}{2}} |\phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2)\cdots\phi_N(\boldsymbol{r}_N)\rangle$$

con esta función se toma el valor esperado del Hamiltoniano

$$E = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$$

Al considerar un conjunto de funciones orbitales $\{\phi_i\}$ ortogonales, esto es

$$\langle\phi_i|\phi_j\rangle = \delta_{ij}$$

y de las propiedades de los determinantes se tiene que

$$E = 2\sum_{i=1}^{N}\langle\phi_i(1)|h(1)|\phi_i(1)\rangle + \sum_{i=1}^{N}\sum_{j=1}^{N}\left[2\left\langle\phi_i(1)\phi_j(2)\left|\frac{1}{r_{12}}\right|\phi_i(1)\phi_j(2)\right\rangle - \left\langle\phi_i(1)\phi_j(2)\left|\frac{1}{r_{12}}\right|\phi_j(1)\phi_i(2)\right\rangle\right]$$

Tomando la variación a primer orden de la energía con respecto a las funciones φ_i y haciendo uso de la teoría de los multiplicadores de Lagrange, la cual permite encontrar máximos y mínimos sujetos a restricciones, se encuentran las ecuaciones que deben satisfacer las funciones espín-orbitales que minimizan el valor de la energía y que además sean ortonormales

$$\delta E + \sum \lambda_{ij}\, \delta_{ij} = 0$$

$$\delta \Bigg\{ 2 \sum_{i=1}^{N} \langle \phi_i(1) | h(1) | \phi_i(1) \rangle + \sum_{i=1}^{N} \sum_{j=1}^{N} \left[2 \left\langle \phi_i(1)\phi_j(2) \middle| \frac{1}{r_{12}} \middle| \phi_i(1)\phi_j(2) \right\rangle - \left\langle \phi_i(1)\phi_j(2) \middle| \frac{1}{r_{12}} \middle| \phi_j(1)\phi_i(2) \right\rangle \right] \Bigg\} + \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{ij} \langle \phi_i | \phi_j \rangle = 0$$

De donde se llega a las ecuaciones de Hartree-Fock

$$h(1)\phi_i(1) + \left[\sum_{j=1}^{N} \int \frac{\phi_j^*(2)\phi_j(2)}{r_{12}} d\boldsymbol{r}_2 \right] \phi_i(1) - \left[\sum_{j=1}^{N} \int \frac{\phi_j^*(2)\phi_i(2)}{r_{12}} d\boldsymbol{r}_2 \right] \phi_j(1) = \epsilon_i \phi_i(1)$$

para $(i = 1,2, \dots, N)$.

Si consideramos que el conjunto de funciones $\{\phi_i\}$ es un conjunto ortonormal, se tiene que estas ecuaciones se pueden escribir como

$$h(1)\phi_i(1) + \left[\sum_{j=1}^{N} \int \frac{\phi_j^*(2)\phi_j(2)}{r_{12}} d\boldsymbol{r}_2 \right] \phi_i(1) - \left[\sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \int \frac{\phi_j^*(2)\phi_i(2)}{r_{12}} d\boldsymbol{r}_2 \right] \phi_j(1) = \epsilon_i \phi_i(1) \qquad \text{para } (i = 1,2, \dots, N)$$

Luego, con el propósito de facilitar el análisis de esta ecuación, se multiplica y divide el término de intercambio por $\phi_i^*(1)\phi_i(1)$, de lo cual se obtiene:

$$h(1)\phi_i(1) + \left[\sum_{j=1}^{N} \int \frac{\phi_j^*(2)\phi_j(2)}{r_{12}} d\boldsymbol{r}_2 \right] \phi_i(1) - \left[\sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \int \frac{\phi_i^*(1)\phi_j^*(2)\left\{\frac{1}{r_{12}}\right\}\phi_i(2)\phi_i(1)}{\phi_i^*(1)\phi_i(1)} d\boldsymbol{r}_2 \right] \phi_j(1) = \epsilon_i \phi_i(1) \qquad \text{para } (i = 1,2, \dots, N)$$

Esto es

$$h(1)\phi_i(1) + \left[\sum_{j=1}^{N} \int \frac{\phi_j^*(2)\phi_j(2)}{r_{12}} d\boldsymbol{r}_2\right]\phi_i(1)$$
$$- \left[\int \frac{\sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \frac{\phi_i^*(1)\phi_j^*(2)\phi_j(1)\phi_i(2)}{\phi_i^*(1)\phi_i(1)}}{r_{12}} d\boldsymbol{r}_2\right]\phi_i(1)$$
$$= \epsilon_i \phi_i(1) \qquad \text{para } (i = 1,2,\ldots,N)$$

Esta ecuación muestra que la función ϕ_i es solución de una ecuación de Schrödinger con un operador hamiltoniano, el cual es la suma de los operadores de energía cinética y energía potencial en el campo de los núcleos. La energía potencial se debe al campo de los N electrones distribuidos en los orbitales ocupados en el determinante de Slater, menos un término que involucra al operador de intercambio; este último potencial corresponde a la distribución de carga.

$$\rho(2) = \sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \frac{\phi_i^*(1)\phi_j^*(2)\phi_j(1)\phi_i(2)}{\phi_i^*(1)\phi_i(1)}$$

A esta densidad se le conoce como densidad de carga de intercambio, y a esta solo contribuyen los electrones que tienen igual espín que el electrón descrito por el orbital ϕ_i, integrando se tiene que

$$\int \rho\,(2) d\boldsymbol{r}_2 = \sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \int \frac{\phi_i^*(1)\phi_j^*(2)\phi_j(1)\phi_i(2)}{\phi_i^*(1)\phi_i(1)} d\boldsymbol{r}_2$$

$$= \frac{\sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right)}{\phi_i^*(1)\phi_i(1)} \phi_i^*(1)\phi_j(1) \int \phi_j^*(2)\phi_i(2)\, d\boldsymbol{r}_2 = 1$$

Con este resultado llegamos a que un electrón descrito por las ecuaciones de Hartree-Fock se mueve en un potencial promedio debido a los *N*-1 electrones restantes. Otra propiedad importante de la densidad de carga de intercambio es reducir la densidad de carga del término de Coulomb en el punto de observación $\boldsymbol{r}$ en una cantidad

$$\sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \phi_j^*(\boldsymbol{r}_1)\phi_j(\boldsymbol{r}_1)$$

Dada la expresión anterior de $\rho(2)$, para $\boldsymbol{r}_2 = \boldsymbol{r}_1$ se tiene que

$$\rho(\boldsymbol{r}_1) = \sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \frac{\phi_i^*(1)\phi_j^*(1)\phi_j(1)\phi_i(1)}{\phi_i^*(1)\phi_i(1)}$$

$$\therefore \qquad \rho(\boldsymbol{r}_1) = \sum_{j=1}^{N} \delta\left(m_s^i, m_s^j\right) \phi_j^*(1)\phi_j(1)$$

Corrección electrónica

La solución obtenida usando únicamente el método de Hartree-Fock no es la esperada de acuerdo con las mediciones experimentales, de modo que para poder corregir el error en esta aproximación se hace lo que se menciona en las siguientes secciones.

Correlación electrónica

Aun cuando la Φ_0 pueda ser una buena solución al problema molecular considerado, no es la función de onda exacta. El método de Hartree-Fock depende de promedios: este no considera las interacciones electrostáticas instantáneas entre los electrones, no toma en cuenta los efectos mecánico-cuánticos sobre las distribuciones de electrones, porque el efecto de *N*-1 electrones sobre un electrón de interés se trata en forma promedio. Una medida de la aproximación de las soluciones que proporciona el método de HF a la descripción de un sistema molecular es la llamada energía de correlación, que Löwdin (1968) define como: "La diferencia de energía de correlación para un cierto estado con respecto a un hamiltoniano es la diferencia entre el valor propio del hamiltoniano exacto y su valor esperado en la aproximación de Hartree-Fock, para el estado en consideración".

En forma matemática viene dado por:

$$E_{corr} = E - E(HF).$$

El método de HF proporciona un número grande de resultados valiosos acerca de la estructura electrónica de las moléculas. Sin embargo, es una aproximación y no incluye correlación electrónica. Existe una buena cantidad de trabajo en el campo del cálculo de la estructura

electrónica que se dirige a tomar en cuenta la correlación electrónica (Atkins y Friedman, 2005).

Configuración de funciones de estado

El método de Hartree-Fock proporciona un conjunto finito de espín-orbitales cuando se lleva a cabo una expansión en un conjunto base finito. En general, una base con *M* miembros resulta en 2*M* diferentes espín-orbitales. Se ordenan los espín-orbitales de acuerdo con su energía, ocupados por los *N* electrones y se forma la función de onda de Hartree-Fock, Φ_0. Ahí permanecen los determinantes de Slater 2*M*-*N* orbitales virtuales. Claramente muchos determinantes de Slater pueden formarse de los 2*M* espín-orbitales, Φ_0 es solo uno de ellos. Para clasificar los demás determinantes de acuerdo con cuantos electrones han sido movidos desde orbitales ocupados a orbitales virtuales se usa la función de onda del determinante de Φ_0 como una referencia, se simplifica la apariencia de las siguientes expresiones omitiendo los factores de normalización de los determinantes de Slater y se reescribe Φ_0 como:

$$\Phi_0 = |\phi_1 \phi_2 \cdots \phi_a \phi_b \cdots \phi_N|$$

donde ϕ_a y ϕ_b están entre los *N* espín-orbitales ocupados para el estado base HF. Un determinante simplemente excitado corresponde a uno para el cual un electrón en el espín orbital ocupado ϕ_a ha sido movido a un espín-orbital ϕ_p.

$$\Phi_a^p = |\phi_1 \phi_2 \cdots \phi_p \phi_b \cdots \phi_N|$$

Un determinante doblemente excitado es aquel en el cual dos electrones han sido promovidos, uno de ϕ_a a ϕ_p y el otro de ϕ_b a ϕ_q

$$\Phi_{ab}^{pq} = |\phi_1 \phi_2 \cdots \phi_p \phi_q \cdots \phi_N|$$

En forma similar, se pueden construir otros determinantes múltiplemente excitados o también construir funciones multielectrónicas que están constituidas por una combinación lineal de un pequeño número de determinantes, los cuales se seleccionan de acuerdo con las restricciones de que la simetría de la molécula sea correcta, a estas funciones se les denomina una configuración de función de estado (CSF por sus siglas en inglés). Una CSF es una función propia de todos los operadores que conmutan con H. Estas CSF excitadas se pueden tomar para aproximar funciones de onda de estados excitados y se pueden usar en

combinación lineal con ϕ_0 para mejorar la representación de la función de onda del estado base (o de cualquier estado excitado) (Atkins y Friedman, 2005).

Interacción de configuraciones

Aun cuando la ecuación de Hartree-Fock es muy exitosa en muchos casos, tiene sus limitaciones. Un ejemplo en el que predice resultados *cualitativamente* incorrectos es el ordenamiento de los potenciales de ionización de N_2. Además, el método de Hartree-Fock restringido no puede describir la disociación de moléculas en fragmentos de capa abierta (p.e., $H_2 \rightarrow 2H$). El procedimiento HF no-restringido da un esquema cualitativamente correcto de tales disociaciones; sin embargo, las curvas de energía potencial resultantes (o superficies en general) no son exactas.

Por más buena que parezca ser la función Φ_0, no es la función de onda exacta. El método de Hartree-Fock se basa en promedios: no considera las interacciones electrónicas *instantáneas* entre electrones; tampoco toma en cuenta los efectos mecánico-cuánticos en la distribución de electrones porque el efecto de N-1 electrones en un electrón de interés se trata de manera promedio. Resumimos estas deficiencias diciendo que el método HF ignora la correlación electrónica.

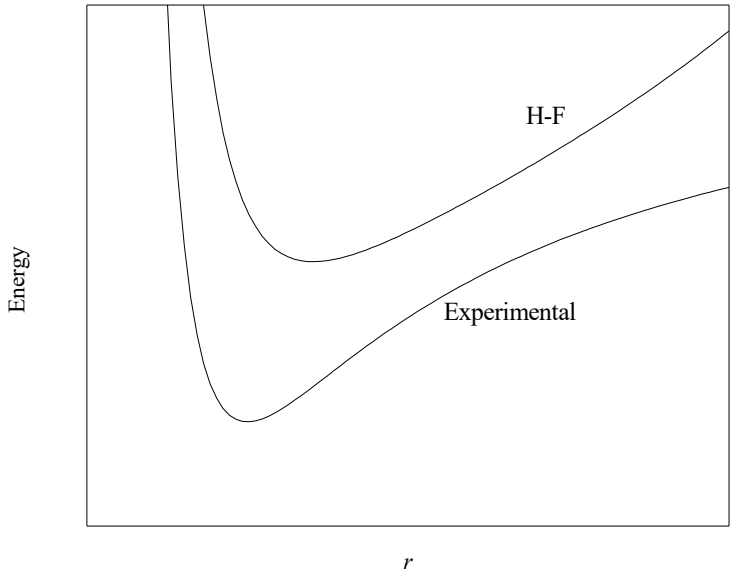
Un método para mejorar la aproximación de Hartree-Fock es precisamente el de *interacción de configuraciones* (CI, por sus siglas en inglés) para obtener la *energía de correlación*. La energía de correlación (E_{corr}) está definida como la diferencia entre la energía no-relativista exacta de la energía del estado base del sistema en la aproximación Born-Oppenheimer ($\mathcal{E}_0$) y la energía de Hartree-Fock (E_0) obtenida en el límite en que el conjunto de bases tiende a ser completo

$$E_{\text{corr}} = \mathcal{E}_0 - E_0$$

Debido a que la energía de Hartree-Fock es una cota superior (figura 31) a la energía exacta, la energía de correlación es negativa. Por otra parte, las curvas de Hartree-Fock frecuentemente se cruzan; sin embargo, la *regla de cruce evitado* dice que estados de moléculas diatómicas de simetría idéntica no se pueden cruzar (ver figuras 31 y 32). Si tenemos dos estados de la misma simetría con funciones de onda ψ_1 y ψ_2 entonces podemos escribir

$$\psi = A_1\psi_1 + A_2\psi_2$$

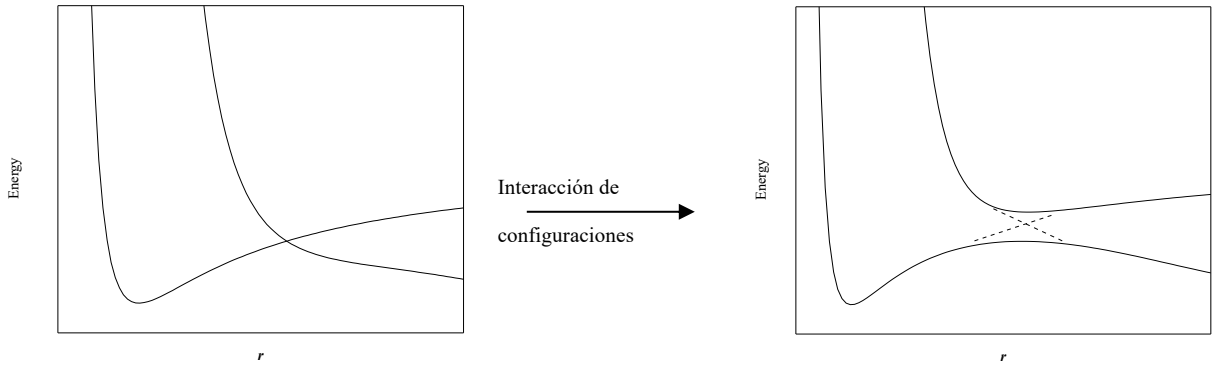
Figura 31. *Comparación entre el método H-F y el experimental*



Fuente: elaboración propia.

donde A_1 y A_2 son coeficientes por encontrarse usando minimización de la energía (Richards y Cooper, 1983). Esta configuración es mezclada y produce curvas de potencial que no violan la regla de no-cruzamiento (como en la figura 32).

Figura 32. *Cruce evitado entre dos curvas de la misma especie*



Fuente: Richards y Cooper (1983).

El procedimiento de Hartree-Fock produce un conjunto $\{\chi_i\}$ de 2K espín-orbitales. El estado base Hartree-Fock

$$\left|\Psi_0\right\rangle = \left|\chi_1\chi_2\cdots\chi_a\chi_b\cdots\chi_N\right\rangle \tag{86}$$

es la mejor (en sentido variacional) aproximación al estado base, de la forma de un solo determinante. Sin embargo, es solo uno de los muchos determinantes que se pueden formar de los 2K>N espín-orbitales. El número de combinaciones de 2K objetos tomados N veces uno a la vez es el coeficiente binomial

$$\binom{2K}{N} = \frac{(2K)!}{N!(2K-N)!}$$

Por ejemplo, si sustituimos N=10, *i.e.*, un sistema con 10 electrones y K=20, *i.e.*, un sistema con 20 conjuntos de funciones base, el número de determinantes considerado es

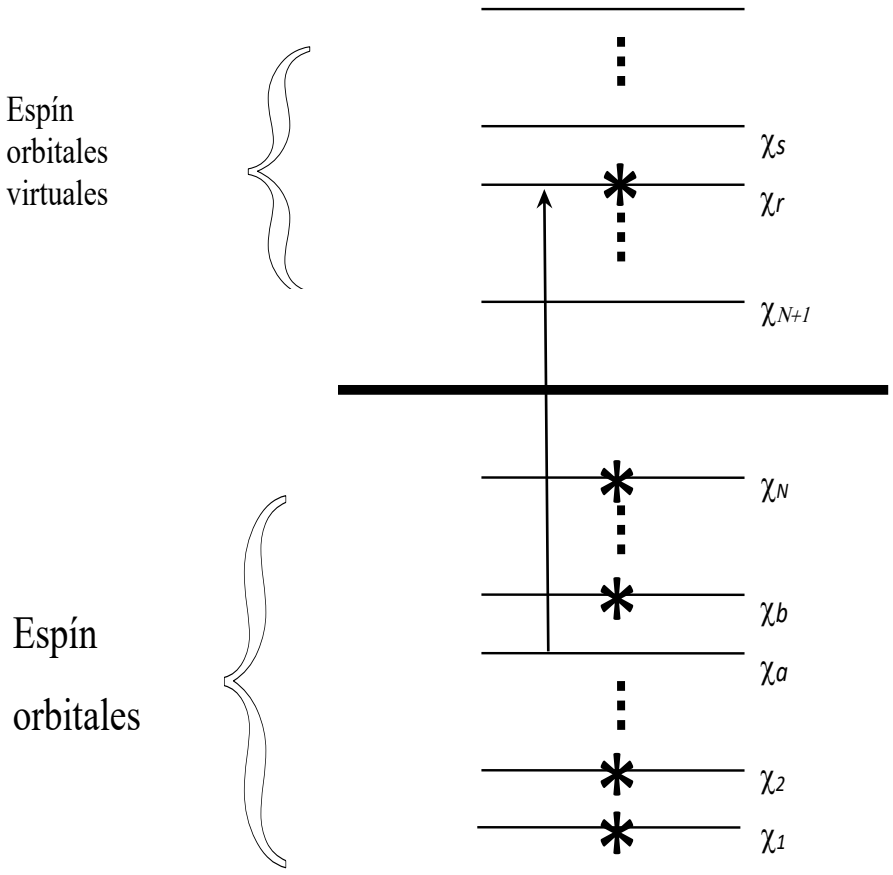
$$\binom{2K}{N} \cong 8.47...\times 10^8$$

El número de combinaciones es lo mismo que el número de determinantes que uno forma de N electrones de 2K espín-orbitales; el estado base Hartree-Fock es solamente uno de estos.

Un determinante simplemente excitado es aquel en el cual un electrón que ocupaba $\{\chi_a\}$ en el estado base Hartree-Fock (86) es promovido a un espín-orbital $\{\chi_r\}$, como en la figura 33:

$$\left|\Psi_a^r\right\rangle = \left|\chi_1\chi_2\cdots\chi_r\chi_b\cdots\chi_N\right\rangle$$

Figura 33. *Salto de un electrón de un orbital a un orbital virtual, corresponde a simple excitación*

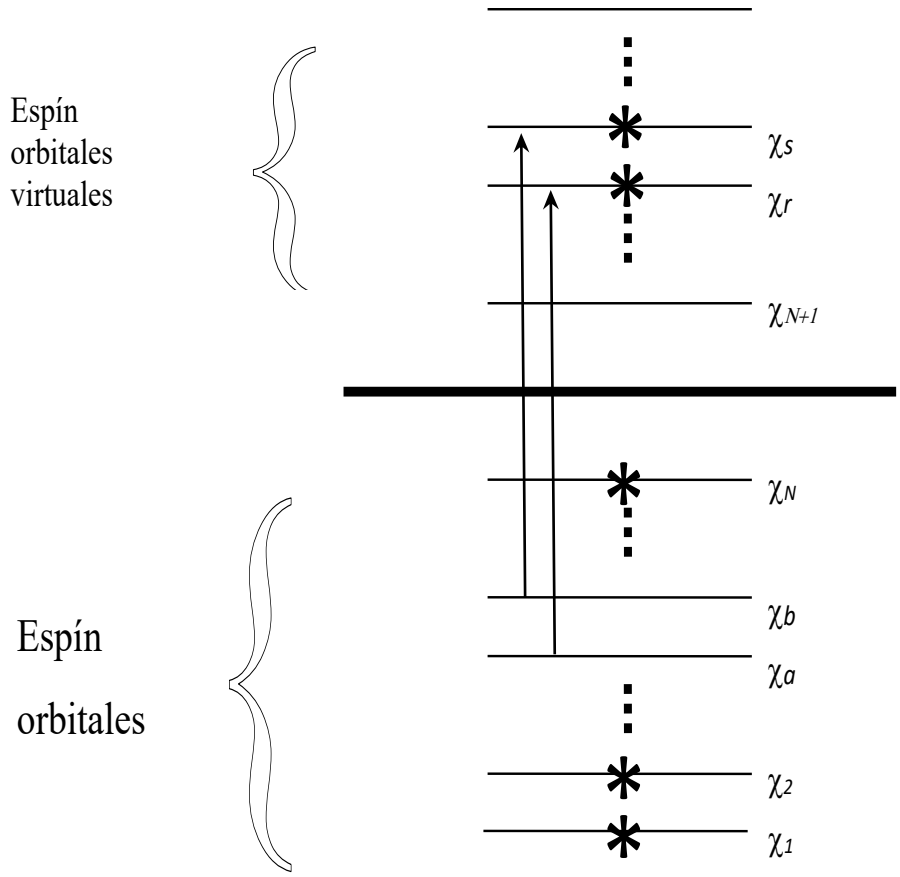


Fuente: Szabo y Ostlund (1996).

Un determinante excitado doblemente, como el que se muestra en la figura 34, es aquel en el cual los electrones son excitados de $\{\chi_a\}$ y $\{\chi_b\}$ a $\{\chi_r\}$ y $\{\chi_s\}$, respectivamente.

$$\left|\Psi_{ab}^{rs}\right\rangle = \left|\chi_1\chi_2\cdots\chi_r\chi_s\cdots\chi_N\right\rangle$$

Figura 34. *Salto de dos electrones de orbitales a orbitales virtuales, corresponde a doble excitación*



Fuente: Szabo y Ostlund (1996).

La idea básica del método de CI es diagonalizar el hamiltoniano N-electrónico en una base de funciones N-electrón (determinantes de Slater). En otras palabras, representamos la función de onda exacta como una combinación lineal de funciones N-electrón de prueba y usamos el método variacional lineal. Si las bases fueran completas, obtendríamos las energías exactas no solo del estado base sino también de todos los estados excitados del sistema.

En principio CI proporciona una solución exacta del problema de muchos electrones. En la práctica, sin embargo, podemos manipular solamente un conjunto finito de funciones de prueba N-electrón; consecuentemente, CI proporciona solo una cota superior a las energías exactas. Dado algún conjunto arbitrario de 2K espín orbitales de un electrón, podemos

construir $\binom{2K}{N}$ diferentes determinantes de Slater N-electrón. Desafortunadamente aún para moléculas pequeñas y conjuntos de bases de un-electrón de tamaño moderado, el número de determinantes N-electrón es verdaderamente enorme. Entonces, aún si uno usa bases de un-electrón finitas, uno usualmente debe truncar la función de prueba de alguna manera y usar solamente una fracción de todas las posibles funciones N-electrón.

Por simplicidad supongamos que nuestra molécula de interés tiene un número par de electrones que a primera aproximación se representa adecuadamente por un determinante HF restringido de capa-cerrada $|\Phi_0\rangle$.

Supongamos que hemos resuelto las ecuaciones de Roothaan en un conjunto de bases finita y obtuvimos un conjunto de 2K espín-orbitales $\{\chi_i\}$ El determinante formado de los N espín-orbitales de energía más baja es $|\Phi_0\rangle$. Además de $|\Phi_0\rangle$ podemos formar un gran número de otros determinantes n-electrón de los 2K espín-orbitales. Es conveniente describir estos determinantes diciendo como difieren de $|\Phi_0\rangle$. Así, el conjunto de posibles determinantes incluyen $|\Phi_0\rangle$, los determinantes simplemente excitados $|\Phi_a^r\rangle$, los determinantes doblemente excitados $|\Phi_{ab}^{rs}\rangle$, etc., hasta incluir determinantes N-veces excitados. Podemos usar estas funciones de onda de muchos electrones como una base en la cual expandir las funciones de onda exactas $|\Psi_0\rangle$ de muchos electrones. Si $|\Phi_0\rangle$ es una aproximación razonable a $|\Psi_0\rangle$, entonces sabemos del principio variacional que una mejor aproximación (la cual llega a ser exacta conforme las bases tiendan a ser completas) es

$$|\Psi_0\rangle = c_0|\Phi_0\rangle + \sum_{ar} c_a^r|\Phi_a^r\rangle + \sum_{\substack{a<b\\r<s}} c_{ab}^{rs}|\Phi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c\\r<s<t}} c_{abc}^{rst}|\Phi_{abc}^{rst}\rangle + \sum_{\substack{a<b<c<d\\r<s<t<u}} c_{abcd}^{rstu}|\Phi_{abcd}^{rstu}\rangle + \cdots \qquad (87)$$

Esta es la forma de la función de onda completa CI. Las restricciones en los índices de las sumatorias (es decir, $a<b$, $r<s$, etc.) aseguran que un determinante excitado dado está incluido en la suma sólo una vez. Cuando se hacen algunas manipulaciones formales, algunas veces es conveniente quitar esta restricción y reescribir la ecuación (87) como

$$|\Psi_0\rangle = c_0|\Phi_0\rangle + \left(\frac{1}{1!}\right)^2 \sum_{ar} c_a^r|\Phi_a^r\rangle + \left(\frac{1}{2!}\right)^2 \sum_{abrs} c_{ab}^{rs}|\Phi_{ab}^{rs}\rangle + \left(\frac{1}{3!}\right)^2 \sum_{\substack{abc\\rst}} c_{abc}^{rst}|\Phi_{abc}^{rst}\rangle + \left(\frac{1}{4!}\right)^2 \sum_{\substack{abcd\\rstu}} c_{abcd}^{rstu}|\Phi_{abcd}^{rstu}\rangle + \cdots$$

Se incluye un factor $(1/n!)^2$ enfrente de la sumatoria para asegurar que una excitación dada es realmente contada sólo una vez. Por ejemplo, la sumatorias no restringidas para dobles excitaciones incluye los términos siguientes

$$c_{ab}^{rs}\left|\Phi_{ab}^{rs}\right\rangle,\ \ c_{ba}^{rs}\left|\Phi_{ba}^{rs}\right\rangle,\ \ c_{ab}^{sr}\left|\Phi_{ab}^{sr}\right\rangle,\ \ \text{y}\ \ c_{ba}^{sr}\left|\Phi_{ba}^{sr}\right\rangle.$$

Ahora, si requerimos que los coeficientes c_{ab}^{rs} sean antisimétricos con respecto al intercambio de *a* y *b* o de *r* y *s* tal como las funciones de onda, entonces los cuatro términos son iguales. Así el factor de 1/4 asegura que el determinante se cuente solo una vez. La ecuación (87) se reescribe de manera más simple como:

$$\Psi_i = \sum_{i=1}^{K} c_{ij}\Phi_j \tag{88}$$

Si tenemos 2K espín orbitales, N estarán ocupados en $\left|\Phi_0\right\rangle$ y 2K – N estarán desocupados. Podemos escoger *n* espín orbitales de aquellos ocupados en

$$\left|\Phi_0\right\rangle \text{ de } \binom{N}{n}$$

maneras. Similarmente podemos escoger *n* orbitales de los 2K – N orbitales virtuales en

$$\binom{2K-N}{n}$$

maneras. Entonces, el número total de determinantes *n*-veces excitados es

$$\binom{N}{n}\binom{2K-N}{n}$$

Aún para moléculas pequeñas y conjuntos de bases de un-electrón solo de tamaño moderado, el número de determinantes *n*-veces excitado es extremadamente grande para toda *n* excepto 0 y 1. Un número significativo de estos determinantes puede ser eliminado (aunque en la mayoría de los casos no suficiente) explotando el hecho de que no hay funciones de onda mezcladas con espín diferente (*i.e.*, $\left\langle\Phi_i\left|H\right|\Phi_j\right\rangle = 0\ \ si\ \ \left|\Phi_i\right\rangle$ y $\left|\Phi_j\right\rangle$ tienen espín diferente).

Dada una función de prueba de la ecuación (87) podemos encontrar las energías correspondientes usando el método variacional lineal. Esto consiste en formar la matriz de

representación del hamiltoniano en la base de las funciones N-electrón de la expansión (87) y entonces encontrar los eigenvalores de esta matriz. Esta es llamada la matriz CI total y el método es referido como CI total. Los eigenvalores más bajos serán una cota superior a la energía del estado base del sistema. Los eigenvalores más grandes serán cotas superiores a estados excitados del sistema. La diferencia entre el eigenvalor más bajo ($\mathcal{E}_0$) y la energía de Hartree-Fock (E_0) obtenida en la misma base de un electrón es llamada la *energía de correlación del conjunto de bases*. Conforme el conjunto de bases de un electrón tiende a la completez, la energía de correlación del conjunto de bases tiende a la energía de correlación exacta. La energía de correlación del conjunto de bases obtenida al realizar un CI total es, sin embargo, exacta en el subespacio abarcado por las bases de un-electrón. Para un conjunto de bases de un-electrón dado, CI total es lo mejor que uno puede hacer.

Métodos de multiconfiguración y multireferencia

En el método de multiconfiguración de campo autoconsistente (MCSCF) se optimizan los coeficientes c_{ij} en la ecuación (88) y los coeficientes de expansión C_{Js} en la ecuación (81). Al optimizar al mismo tiempo los dos conjuntos de coeficientes de expansión se demanda mucho tiempo de computadora, aunque optimizando los C_{Js} se pueden obtener resultados más exactos con un número relativamente pequeño de CSF. A cada determinante o una combinación lineal de un pequeño número de ellos construidos para tener una simetría electrónica correcta se le llama una función de configuración de estado (CSF).

El desarrollo de los métodos eficientes MCSCF es activamente investigado, y es muy importante para los estados excitados. Uno de tales esquemas es el método de campo autoconsistente de espacio activo completo (CASSCF). En este enfoque los espín-orbitales (que también son optimizados durante el cálculo para determinar los valores de los C_{Js}) se dividen en tres clases:

a) Un conjunto de orbitales inactivos compuestos de los espín-orbitales de energía más baja que son doblemente ocupados en todos los determinantes.
b) Un conjunto de espín-orbitales virtuales de muy alta energía que estén desocupados en todos los determinantes

c) Un conjunto de orbitales activos que sean energéticamente intermedios entre los orbitales inactivos y los virtuales.

Los electrones activos son los electrones que no están en el conjunto de orbitales inactivos doblemente ocupados. Los CSF incluidos en el cálculo CASSCF son configuraciones (de la simetría y espín apropiados) que se originan de todas las formas posibles de la distribución de los electrones activos sobre los orbitales activos.

Hay tendencias para escoger cuales orbitales incluir como orbitales activos. Una tendencia consiste en tomar los orbitales activos como aquellos orbitales (enlazantes, no enlazantes y antienlazantes) que se originan en la teoría de orbitales moleculares MO, de los orbitales atómicos de valencia que conforman a la molécula (Atkins y Friedman, 2005).

En los métodos convencionales CI descritos anteriormente, la función de onda SCF de Hartree-Fock Ψ_0 se usa como una configuración de referencia y las funciones del estado de configuración se forman sacando electrones de los espín-orbitales ocupados de Ψ_0 para llevarlos hacia espín-orbitales desocupados. En la interacción de configuración de multireferencia (MRCI) se origina un conjunto de configuraciones de referencia, a partir del cual se crean los determinantes excitados para su uso en un cálculo CI.

Como ejemplo de lo anterior, se puede hacer un cálculo MCSCF y tener un conjunto de configuraciones de referencia compuesto de aquellos determinantes con un coeficiente c_{ij} mayor que algún valor predeterminado (como 0.05 o 0.1) (Atkins y Friedman, 2005) en la función de onda final normalizada MCSCF. Para cada uno de los determinantes de referencia los electrones se mueven de espín-orbitales ocupados a espín-orbitales desocupados para crear más determinantes que sean adicionados en la expansión CI de la ecuación (88). Posteriormente se realiza el CI, optimizando todos los coeficientes c_{ij} de los determinantes que han sido incluidos. En los determinantes de referencia, frecuentemente determinantes simple y doblemente excitados con respecto a Ψ_0, se incluyen excitaciones simples y dobles de los determinantes de referencia. Por lo tanto, la función de onda MRCI final incluye determinantes que son triple y cuádruplemente excitados de Ψ_0. Debido a que los cálculos MRCI frecuentemente contienen los determinantes cuádruplemente excitados, estos determinantes usualmente reducen significativamente el error de consistencia de tamaño encontrado en cálculos CISD.

Usualmente es necesario que la energía de un sistema de *N* partículas aún en presencia de interacciones sea proporcional al número de estas partículas, lo cual se conoce como la consistencia de tamaño. En particular, la energía del sistema AB calculado cuando los subsistemas A y B están muy separados debe ser igual a la suma de las energías de A y B por separado calculadas usando el mismo método. Esta es una gran deficiencia del método de cálculos CI limitado.

Frecuentemente una fracción grande de la energía de correlación exacta puede recuperarse de cálculos MRCI con un número mucho menor de determinantes.

Teoría de perturbación de muchos cuerpos de Möller-Plesset

Los cálculos de interacción de configuración CI proporcionan una aproximación sistemática para ir a un nivel de energía de Hartree-Fock más profundo, debido a la inclusión de determinantes que son excitados simplemente, doblemente, triplemente y así sucesivamente, de una configuración de referencia. Una característica importante del método es que es variacional, pero su desventaja es que carece de consistencia de tamaño (exceptuando el CI completo). La teoría de perturbación (PT) proporciona un acercamiento sistemático alternativo para encontrar la energía de correlación, además, sus cálculos son consistentes en tamaño, no son variacionales en el sentido de que, en general, no dan las energías que son límites superiores a la energía exacta. La aplicación de PT a un sistema compuesto de muchas partículas interactuantes generalmente se llama teoría de perturbación de muchos cuerpos (MBPT) (Pople et al., 1976). Como se requiere encontrar la energía de correlación para el estado base se toma el hamiltoniano de orden cero de los operadores de Fock del método SCF de Hartree-Fock.

En la teoría de perturbación el hamiltoniano real H, y el hamiltoniano del modelo más simple del sistema $H^{(0)}$, difieren por una contribución:

$$H = H^{(0)} + H^{(1)}$$

donde $H^{(1)}$ es la perturbación.

Möller y Plesset (1934) han hecho una selección de $H^{(0)}$ llamada teoría de perturbación de Möller-Plesset (MPPT). Las aplicaciones de MPPT a sistemas moleculares no comenzaron sino hasta 40 años más tarde.

El hamiltoniano de orden cero $H^{(0)}$, denotado como H_{HF} en este contexto, se da por la suma de operadores de Fock de un electrón

$$H_{HF} = \sum_{i=1}^{n} f_i$$

La función de onda Ψ_0 del estado base HF es una eigenfunción de H_{HF} con un eigenvalor $E^{(0)}$ dado por la suma de las energías orbitales de todos los espín-orbitales ocupados. La perturbación $H^{(1)}$ corresponde a

$$H^{(1)} = H - \sum_{i=1}^{n} f_i \,.$$

En la ecuación anterior, H es el hamiltoniano electrónico. La energía HF, E_{HF}, la cual está asociada con la función de onda normalizada Ψ_0 del estado base HF es el valor esperado:

$$E_{HF} = \left\langle \Psi_0 \middle| H \middle| \Psi_0 \right\rangle$$

lo que equivale a

$$E_{HF} = \left\langle \Psi_0 \middle| H_{HF} + H^{(1)} \middle| \Psi_0 \right\rangle \tag{89}$$

Es fácil demostrar que E_{HF} es igual a la suma de la energía de orden cero $E^{(0)}$ y de la corrección a la energía de primer orden $E^{(1)}$. Es decir, dado que Ψ_0 es una eigenfunción de H_{HF} se tiene que

$$E^{(0)} = \left\langle \Psi_0 \middle| H_{HF} \middle| \Psi_0 \right\rangle \tag{90}$$

y que

$$E^{(1)} = \left\langle \Psi_0 \middle| H^{(1)} \middle| \Psi_0 \right\rangle \tag{91}$$

Entonces, se concluye de las ecuaciones (87-89) que

$$E_{HF} = E^{(0)} + E^{(1)}$$

Por lo tanto, la primera corrección a la energía del estado base está dada por la teoría de perturbación a segundo orden como:

$$E^{(2)} = \sum_{J \neq 0} \frac{\left\langle \Psi_J \middle| H^{(1)} \middle| \Psi_0 \right\rangle \left\langle \Psi_J \middle| H^{(1)} \middle| \Psi_0 \right\rangle}{E^{(0)} - E_J}$$

Para evaluar esta ecuación, se calculan los elementos de matriz fuera de la diagonal $\left\langle \Psi_J \left| H^{(1)} \right| \Psi_J \right\rangle$. Primero, nótese que el elemento de matriz

$$\left\langle \Psi_J \left| H_{HF} \right| \Psi_0 \right\rangle = 0$$

dado que Ψ_0 es una eigenfunción de H_{HF} y de los espín-orbitales, entonces los determinantes son ortogonales. Por lo tanto, si $\left\langle \Psi_J \left| H \right| \Psi_0 \right\rangle = 0$, entonces $\left\langle \Psi_J \left| H^{(1)} \right| \Psi_0 \right\rangle = 0$.

Se concluye del teorema de Brillouin (este teorema dice que determinantes simplemente excitados $\left| \Psi_a^r \right\rangle$ no interaccionarán directamente con un determinante Hartree-Fock de referencia $\left| \Psi_0 \right\rangle$, (*i.e.*, $\left\langle \Psi_0 \left| H \right| \Psi_a^r \right\rangle = 0$) que solo los determinantes doblemente excitados tienen elementos de matriz $H^{(1)}$ diferentes de cero con Ψ_0 y, por lo tanto, únicamente las excitaciones dobles contribuyen a $E^{(2)}$. La siguiente expresión es resultado de un análisis de estos elementos de matriz no nulos

$$E^{(2)} = \frac{1}{4} \sum_{a,b}^{ocu} \sum_{r,s}^{vir} \frac{(ab \| rs)(rs \| ab)}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s},$$

donde:

$$(ab \| rs) = \int \chi_a^*(1) \chi_b^*(2) \frac{1}{r_{12}} \chi_r(1) \chi_s(2) d\vec{x}_1 d\vec{x}_2$$

$$- \int \chi_a^*(1) \chi_b^*(2) \frac{1}{r_{12}} \chi_s(1) \chi_r(2) d\vec{x}_1 d\vec{x}_2 ,$$

con Ψ_a y Ψ_b espín-orbitales ocupados, con Ψ_r y Ψ_s espín-orbitales virtuales. La anexión de la corrección a la energía a segundo orden se designa como MP2. Se puede extender el MPPT para anexar correcciones a la energía de tercero y cuarto orden, conocidos como MP3 y MP4. Conforme se mueve a ordenes mayores de la teoría de perturbación, el álgebra involucrada se hace más y más complicada y es común usar técnicas diagramáticas para clasificar y representar los diversos términos que aparecen en las expresiones en serie de perturbación, y se usan para probar que MPPT es consistente en tamaño en todos los órdenes.

Las trayectorias de energía mínima (MEP) correspondientes a cada uno de los estados electrónicos se determinan en cada punto mediante cálculos *ab initio* Hartree-Fock de campo

autoconsistente (HF-SCF) con el programa PSHF, usando potenciales relativistas de carozo efectivo (RECP). Estos cálculos son seguidos por cálculos multiconfiguracionales de campo autoconsistente (MCSCF), el cual provee los conjuntos óptimos de orbitales moleculares e integrales para usarse en el esquema de (CI).

Los efectos de correlación electrónica se toman en cuenta usando una interacción de configuración multirreferencial con una función de onda multiconfiguracional de orden cero para perturbación a través del algoritmo iterativo seleccionado para la interacción de configuración (CIPSI) en su versión clase dos. La energía total que provee CIPSI es la suma de la energía variacional en el espacio de referencia y la contribución a la energía de segundo orden Möller-Plesset (MP2) de los determinantes fuera del espacio de referencia.

Teoría de funcionales de la densidad

Las dificultades matemáticas en la solución numérica de la ecuación de Schrödinger por el método de Hartree-Fock incrementan enormemente conforme aumenta el número de electrones en un átomo. Esto hace que este método sea escasamente aplicado al considerar átomos pesados, o moléculas. El método de Thomas-Fermi originalmente se introdujo para calcular la densidad de una distribución de electrones de alguno de los átomos pesados.

La idea original del modelo de Thomas-Fermi

Las hipótesis argumentadas por Thomas son

1) Los electrones están distribuidos uniformemente en el espacio fase seis-dimensional para el movimiento de un electrón en la razón de 2 de cada h^3 de volumen.
2) Existe un potencial efectivo de campo que se determina así mismo por la carga nuclear y su distribución de electrones

Estas son consideraciones estadísticas que se pueden usar para aproximar la distribución de electrones en un átomo. La fórmula de Thomas-Fermi para la densidad electrónica se puede derivar de estas suposiciones (Davydov, 1965). Una derivación distinta se deduce al resolver el problema de una partícula en una caja de potencial usando directamente la ecuación de Schrödinger independiente del tiempo.

Partícula en una caja

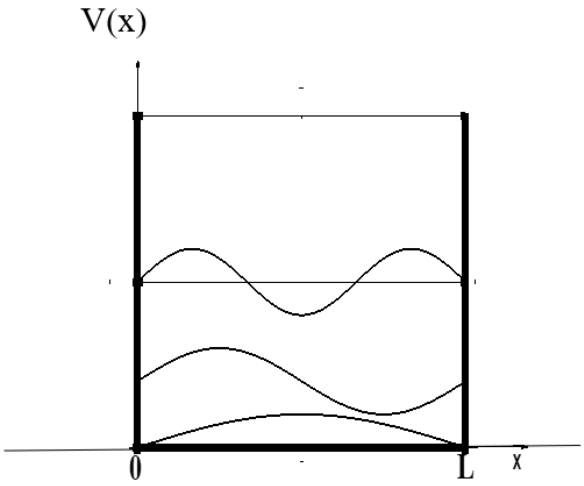
En el caso unidimensional de una partícula en una caja, el hamiltoniano del sistema es

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$$

$$V(x) = \begin{cases} 0 & para\ 0 \le x \le L \\ \infty & de\ otro\ modo \end{cases}$$

Como en la figura 35:

Figura 35. *Caja de potencial unidimensional*



Fuente: elaboración propia.

La ecuación electrónica de Schrödinger en la caja de potencial de la figura 35 es

$$H\psi = E\psi$$

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi = E\psi$$

Para $V(x) = 0$ en $0 \le x \le L$ se tiene

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi = E\psi \implies \frac{d^2\psi}{dx^2} + k^2\psi = 0\ \ con\ k^2 = \frac{2mE}{\hbar^2}\ \ o\ bien\ \ k\hbar = \sqrt{2mE}$$

Cuya ecuación característica es

$$r^2 + k^2 = 0 \qquad \therefore \qquad r = \pm k$$

$$\psi(x) = Ae^{kx} + Be^{-kx}$$

Como $e^{\pm i\theta} = cos\theta \pm isen\theta$ (fórmula de Euler), entonces

$$\psi(x) = A(coskx + isenkx) + B(coskx - isenkx)$$

$$\psi(x) = Acoskx + Aisenkx + Bcoskx - Bisenkx$$

$$\psi(x) = (A + B)coskx + (A - B)isenkx$$

$$\psi(x) = Ccoskx + Dsenkx \quad con \quad C = (A + B) \;\; y \;\; D = i(A - B)$$

Aplicando condiciones a la frontera:

Primero, en $x = 0, se\ tiene\ que\ cos0 = 1,\ sen0 = 0$ por lo tanto $\psi(0) = C$.

Además, las funciones de onda son cero afuera de la caja donde $x < 0\ \ o\ \ x > L$. Las funciones de onda son continuas en todas partes. Esto es que las funciones de onda deben ser cero en las paredes. Por lo tanto, las condiciones de frontera son

$$\psi(0) = 0 \quad y \quad \psi(L) = 0$$

Esto implica que $\psi(0) = C = 0$.

Segundo, en $x = L$

$$\psi(L) = CcoskL + DsenkL = 0$$

$$0 + DsenkL = 0$$

Si $D = 0$ se ncumple la condición de frontera, y como la función de onda es cero, entonces no hay partícula en ninguna parte. La única alternativa es que el seno desaparezca, lo cual solo es posible si kL es igual a un múltiplo entero de π, esto es $kL = n\pi$ con $n = 1,2,\ldots,$

$$k = \frac{n\pi}{L} \qquad n = 1,2,\ldots,$$

$n = 0$ se excluye porque resultaría que la partícula no se encuentra en ninguna parte. El entero n es un *número cuántico* que etiqueta un estado del sistema, y que se puede usar para calcular el valor de una observable del sistema, y que de hecho es la energía en este caso, ya que $E = k^2\hbar^2/2m$, y como $\hbar = {}^{h}/_{2\pi}$, entonces

$$E_n = \frac{n^2\hbar^2\pi^2}{2mL^2}\ o\ bien\ E_n = \frac{n^2h^2}{8mL^2} \qquad n = 1,2,\ldots,$$

es decir, la energía está cuantizada.

En el caso del pozo cuadrado tridimensional se tiene que

$$H = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V(x,y,z)$$

$$V(x,y,z)=\begin{cases}0 & para\ 0\le x\le L_1, 0\le y\le L_2, 0\le z\le L_3,\\ \infty & de\ otro\ modo\end{cases}$$

En forma de operadores la ecuación de Schrödinger está dada por $H\psi = E\psi$, entonces

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V(x,y,z)\right]\psi = E\psi$$

Entre las paredes el potencial es cero, y la función de onda se obtiene al resolver la ecuación diferencial parcial

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2} = -\frac{2mE}{\hbar^2}\psi$$

Las condiciones de frontera son aquellas en que la función de onda desaparece en todas las paredes. Resolvemos esta ecuación separando variables; esto es, la solución de prueba se escribe $\psi(x,y,z) = XYZ$, donde X es función solo de x, Y es función solo de y, y Z es función solo de z. Por ello:

$$YZ\frac{\partial^2 X}{\partial x^2}+XZ\frac{\partial^2 Y}{\partial y^2}+XY\frac{\partial^2 Z}{\partial z^2} = -\frac{2mE}{\hbar^2}XYZ$$

Al dividir entre XYZ se obtiene

$$\frac{X''}{X}+\frac{Y''}{Y}+\frac{Z''}{Z} = -\frac{2mE}{\hbar^2}$$

Donde $X'' = \frac{d^2X}{dx^2}, Y'' = \frac{d^2Y}{dy^2}$, y $Z'' = \frac{d^2Z}{dz^2}$. Concluimos que la ecuación original se puede separar en tres partes

$$X'' = -\left(\frac{2mE_x}{\hbar^2}\right)X, \qquad Y'' = -\left(\frac{2mE_y}{\hbar^2}\right)Y, \qquad Z'' = -\left(\frac{2mE_z}{\hbar^2}\right)Z$$

Con $E_x + E_y + E_z = E$. Estas ecuaciones tienen la misma forma que la ecuación de un sistema unidimensional, y las condiciones de frontera son las mismas. Por lo tanto, la solución es

$$\psi_{n_1 n_2 n_3}(x,y,z) = \frac{2}{(L_1 L_2 L_3)^{1/2}} sen\left(\frac{n_1 \pi x}{L_1}\right) sen\left(\frac{n_2 \pi y}{L_2}\right) sen\left(\frac{n_3 \pi z}{L_3}\right)$$

$$E_{n_1 n_2 n_3} = \frac{h^2}{8m}\left(\frac{n_1^2}{L_1^2} + \frac{n_2^2}{L_2^2} + \frac{n_3^2}{L_3^2}\right) \qquad n_1 = 1,2,\dots,; n_2 = 1,2,\dots,; n_3 = 1,2,\dots,$$

Si dividimos el espacio en muchos cubos pequeños (celdas), cada volumen $\Delta V = \ell^3$ de lado ℓ contiene un número fijo de electrones ΔN (que pueden tener diferentes valores para diferentes celdas), y suponemos que los electrones en cada celda se comportan como fermiones independientes, a la temperatura 0 K, con las celdas independientes entre sí. Entonces los niveles de energía de una partícula en un pozo infinito tridimensional están dados por la fórmula

$$\varepsilon(n_x, n_y, n_z) = \frac{h^2}{8m\ell^2}\left(n_x^2 + n_y^2 + n_z^2\right)$$
$$= \frac{h^2}{8m\ell^2} R^2$$

donde $n_x, n_y, n_z = 1,2,3,\dots,$ y la segunda igualdad define la cantidad R. Para números cuánticos grandes, es decir, para R grande, el número de niveles de energía distintos, con energía más pequeña que ε, se puede aproximar por el volumen de un octante de una esfera con radio R en el espacio (n_x, n_y, n_z). Este número es

$$\Phi(\varepsilon) = \frac{1}{8}\left(\frac{4\pi R^3}{3}\right) = \frac{\pi}{6}\left(\frac{8m\ell^2\varepsilon}{h^2}\right)^{3/2}$$

El número de niveles de energía entre $\varepsilon\ y\ \varepsilon + \delta\varepsilon$ es

$$g(\varepsilon)\Delta\varepsilon = \Phi(\varepsilon + \delta\varepsilon) - \Phi(\varepsilon)$$

$$g(\varepsilon)\Delta\varepsilon = \frac{\pi}{4}\left(\frac{8m\ell^2}{h^2}\right)^{3/2} \varepsilon^{1/2}\delta\varepsilon + O\left((\delta\varepsilon)^{1/2}\right)$$

donde la función $g(\varepsilon)$ es la densidad de estados en la energía ε.

Para computar la energía total para la celda con ΔN electrones, necesitamos la probabilidad de que el estado con energía ε este ocupado, a la cual llamaremos $f(\varepsilon)$. Esta es la distribución de Fermi-Dirac,

$$f(\varepsilon) = \frac{1}{1 + e^{\beta(\varepsilon-\mu)}}$$

La cual a 0 K se reduce a una función escalón

$$f(\varepsilon) = \begin{cases} 1, & \varepsilon < \varepsilon_f \\ 0, & \varepsilon > \varepsilon_f \end{cases} \qquad conforme \quad \beta \longrightarrow \infty$$

Donde ε_f es la energía de Fermi. Todos los estados con energía más pequeña que ε_f están ocupados, y aquellos con energía más grande que ε_f están desocupados. La energía de Fermi ε_f es la temperatura-cero límites del potencial químico μ.

La energía total de los electrones en esta celda se encuentra sumando las contribuciones de los diferentes estados de energía:

$$\Delta \mathrm{E} = 2\int \varepsilon f(\varepsilon) g(\varepsilon) d\varepsilon$$

$$= 4\pi \left(\frac{2m}{h^2}\right)^{3/2} \ell^3 \int_0^{\varepsilon_f} \varepsilon^{3/2}\, d\varepsilon$$

$$= \frac{8\pi}{5}\left(\frac{2m}{h^2}\right)^{3/2} \ell^3 \varepsilon_f^{5/2}$$

Donde el factor 2 entra debido a que cada nivel de energía está doblemente ocupado por un electrón con espín α y otro con espín β. La energía de Fermi ε_f está relacionada al número de electrones ΔN en la celda, a través de la fórmula

$$\Delta N = 2\int f(\varepsilon) g(\varepsilon) d\varepsilon$$

$$= \frac{8\pi}{3}\left(\frac{2m}{h^2}\right)^{3/2} \ell^3 \varepsilon_f^{3/2}$$

$$\Longrightarrow \quad \varepsilon_f = \left(\frac{3}{8\pi}\right)^{2/3} \frac{h^2}{2m}\left(\frac{\Delta N}{\ell^3}\right)^{2/3}$$

$$\Delta E = \frac{3}{5}\Delta N \varepsilon_f$$

o solamente en términos de ΔN se tiene que

$$\Delta E = \frac{3}{5}\frac{\ell^3}{\ell^3}\Delta N\left(\frac{\Delta N}{\ell^3}\right)^{2/3}\left(\frac{3}{8\pi}\right)^{2/3}\frac{h^2}{2m}$$
$$= \frac{3h^2}{10m}\left(\frac{3}{8\pi}\right)^{2/3}\ell^3\left(\frac{\Delta N}{\ell^3}\right)^{5/3}$$

Esta ecuación es una relación entre la energía cinética total y la densidad del electrón

$$\rho = \frac{\Delta N}{\ell^3} = \frac{\Delta N}{\Delta V}$$

para cada celda en el espacio. Agregando las contribuciones de todas las celdas, encontramos que la energía cinética es (en unidades atómicas)

$$T_{TF} = C_F \int \rho^{5/3}(\boldsymbol{r})d\boldsymbol{r}$$

donde

$$C_F = \frac{3h^2}{10m}\frac{4\pi^2}{4\pi^2}\frac{e^2}{e^2}\left(\frac{3}{8\pi}\right)^{2/3}$$
$$= \frac{3}{10}a_H e^2\left(\frac{3(4\pi^2)^{3/2}}{8\pi}\right)^{2/3}$$

Haciendo $a_H e^2 = 1$ se obtiene

$$C_F = \frac{3}{10}(3\pi^2)^{2/3} = 2.871$$

donde $\frac{\Delta N}{\Delta V} = \rho(\boldsymbol{r})$ en el límite cuando $\Delta V \longrightarrow 0$. Entonces T_{TF} es la energía cinética funcional de Thomas-Fermi que se aplicó a electrones en átomos, y que conduce a la aproximación de densidad local (LDA, por sus siglas en inglés), de la cual se determinan propiedades electrónicas de sistemas electrónicos homogéneos. Obviamente, el potencial corresponde al de un pozo cuadrado.

La energía como un funcional de la densidad es

$$E_{TF}[\rho(\boldsymbol{r})] = T_{TF}[\rho] + V_{ne}[\rho] + V_{ee}[\rho]$$

$$E_{TF}[\rho(\boldsymbol{r})] = C_F\int \rho^{5/3}(\boldsymbol{r})d\boldsymbol{r} - Z\int\frac{\rho(\boldsymbol{r})}{r}d\boldsymbol{r} + \frac{1}{2}\iint\frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}d\boldsymbol{r}_1 d\boldsymbol{r}_2$$

es la energía de Thomas-Fermi de átomos. Para moléculas se modifica el segundo término apropiadamente. La densidad electrónica del estado base debe satisfacer el principio variacional

$$\delta\left\{E_{TF}[\rho] - \mu_{TF}\left(\int \rho(\boldsymbol{r})d\boldsymbol{r} - N\right)\right\} = 0$$

De la cual el cálculo variacional permite la ecuación de Euler-Lagrange:

$$\mu_{TF} = \frac{\delta E_{TF}[\rho]}{\delta\rho(\boldsymbol{r})} = \frac{5}{3}C_F\rho^{2/3}(\boldsymbol{r}) - \phi(\boldsymbol{r})$$

donde $\phi(\boldsymbol{r})$ es el *potencial electrostático* en el punto $\boldsymbol{r}$ debido al núcleo y a la distribución electrónica total:

$$\phi(\boldsymbol{r}) = \frac{Z}{r} - \int \frac{\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}d\boldsymbol{r}_2$$

El método primitivo descrito se debilita en moléculas, ya que no predice energías de amarre moleculares. Además, la exactitud para átomos no es tan buena como la de otros métodos. Esto restó importancia para predicciones cuantitativas en física atómica o molecular.

Energía como un funcional de la densidad

Al modelo de Thomas-Fermi para la energía como un funcional de la densidad le falta generalidad, que básicamente consiste en que el potencial externo sea un funcional de la densidad. Los teoremas fundamentales que demuestran esto, son los teoremas de Hohenberg y Kohn.

Teoremas de Hohenberg y Kohn

1. El potencial externo $v(\boldsymbol{r})$ se determina por la densidad electrónica $\rho(\boldsymbol{r})$ en una constante aditiva. No pueden existir dos potenciales $v(\boldsymbol{r})$ diferentes entre sí que generen la misma densidad electrónica $\rho(\boldsymbol{r})$.
2. Para una densidad de prueba $\tilde{\rho}(\boldsymbol{r})$, tal que $\tilde{\rho}(\boldsymbol{r}) \geq 0$ y $\int \tilde{\rho}\,(\boldsymbol{r})d\boldsymbol{r} = N$, se tiene que

$$E_0 \leq E_v[\tilde{\rho}]$$

donde $E_v[\tilde{\rho}]$ es la energía funcional

$$E_v[\rho] = T[\rho] + V_{ne}[\rho] + V_{ee}[\rho]$$
$$= \int \tilde{\rho}\,(\boldsymbol{r}) v(\boldsymbol{r}) d\boldsymbol{r} + F_{HK}[\rho]$$

con $F_{HK}[\rho] = T[\rho] + V_{ee}[\rho]$ y

$$V_{ee}[\rho] = J[\rho] + \text{término no} - \text{clásico}$$

el término $J[\rho]$ es la repulsión clásica, y la mayor parte del término no-clásico está definida como *energía de intercambio y correlación*.

Ecuación de Kohn-Sham

El hamiltoniano de un sistema de partículas independientes en un sistema de referencia no-interactuante se puede escribir como

$$\widehat{H}_s = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2\right) + \sum_{i=1}^{N} v_s(\boldsymbol{r}_i)$$

En el cual no hay términos de repulsión electrón-electrón y para el cual la densidad electrónica del estado base es exactamente ρ. Para este sistema hay un determinante exacto de la función de onda del estado base

$$\Psi_s = \frac{1}{\sqrt{N!}} det[\psi_1 \psi_2 \cdots \psi_N]$$

donde las ψ_i son los N estados más bajos del hamiltoniano de un-electrón $\hat{h}_s$

$$\hat{h}_s \psi_i = \left[-\frac{1}{2}\nabla_i^2 + v_s(\boldsymbol{r}_i)\right]\psi_i = \varepsilon_i \psi_i$$

Un determinante de Slater es la función de onda que representa correctamente este sistema y, por lo tanto, la energía cinética es

$$T_s[\rho] = \left\langle \Psi_s \middle| \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2\right) \middle| \Psi_s \right\rangle$$
$$= \sum_{i=1}^{N} \left\langle \psi_i \middle| -\frac{1}{2}\nabla_i^2 \middle| \psi_i \right\rangle$$

y la densidad electrónica se escribe como

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} |\psi_i(\boldsymbol{r})|^2$$

los orbitales que minimizan la energía deben cumplir con

$$\left(-\frac{1}{2}\nabla_i^2 + v_s(\boldsymbol{r}_i)\right)\psi_i(\boldsymbol{r}) = \epsilon_i\psi_i(\boldsymbol{r})$$

Kohn y Sham usan como referencia un sistema de partículas independientes

$$T_s = \sum_{i=1}^{N}\left\langle\chi_i\left|-\frac{1}{2}\nabla_i^2\right|\chi_i\right\rangle,$$

$$\rho_{ref}(\boldsymbol{r}) = \sum_{i=1}^{N}\int d\omega\,\chi_i^*(\boldsymbol{x})\chi_i(\boldsymbol{x}) = \sum_{i=1}^{N}\psi_i^*(\boldsymbol{x})\psi_i(\boldsymbol{x})$$

La energía cinética del sistema de referencia no representa a la energía cinética T del estado base, por lo tanto, existe la diferencia

$$\Delta T = T - T_s$$

La interacción electrón-electrón, del funcional universal de Hohenberg y Kohn, se divide en dos partes:

$$V_{ee}^{coul} = J[\rho] = \frac{1}{2}\iint\frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{r}'d\boldsymbol{r}$$

y una parte no-coulómbica. Con estos elementos, el funcional universal toma la forma

$$F[\rho] = T[\rho] + V_{ee} = \Delta T + T_s + J[\rho] + V_{ee}^{no-coul}$$

y la energía total del sistema E_{KS} es

$$E_{KS}[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int\rho\,(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r}$$

$$= \sum_{i=1}^{s}\sum_{s}\int\psi_i^*(\boldsymbol{r})\left(-\frac{1}{2}\nabla^2\right)\psi_i(\boldsymbol{r})d\boldsymbol{r} + J[\rho] + E_{xc}[\rho] + \int\rho\,(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r}$$

El funcional de intercambio y correlación debe ser de tal forma que la densidad electrónica que minimiza a la energía de Kohn-Sham sea igual a la densidad electrónica del estado base. Los orbitales que minimizan la energía deben satisfacer

$$\left(-\frac{1}{2}\nabla_i^2 + \int\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}' + \frac{\delta E_{xc}}{\delta\rho(\boldsymbol{r})} + v(\boldsymbol{r})\right)\chi_i(\boldsymbol{r}) = \epsilon_i\chi_i(\boldsymbol{r})$$

O bien

$$\left(-\frac{1}{2}\nabla_i^2 + v_{KS}^{efect}(\boldsymbol{r})\right)\chi_i(\boldsymbol{r}) = \epsilon_i\chi_i(\boldsymbol{r})$$

Una forma de aproximar al funcional de intercambio y correlación es suponer que esta cantidad se puede dividir en dos contribuciones, es decir:

$$E_{xc} = E_x + E_c$$

Para el intercambio ya conocemos la expresión de Hartree-Fock, que en el contexto de Kohn-Sham se le conoce como el intercambio exacto. Partiendo de la expresión del intercambio exacto y evaluándola para el gas de electrones (ondas planas) se obtiene

$$E_x = -c_x \int \rho(\boldsymbol{r})^{4/3} d\boldsymbol{r}$$

Mientras que la fórmula de energía de correlación es

$$E_c\left[\rho^{\alpha}, \rho^{\beta}\right] = \int \rho E_c(\rho, \zeta) d\boldsymbol{r}$$

Las fórmulas más exactas disponibles para la energía de correlación por partícula de un gas-uniforme se han obtenido al aplicar un método cuántico de Montecarlo, y un análisis de una aproximación de fase aleatoria. Por lo tanto, esto sigue abierto a soluciones que no necesiten cálculos aleatorios.

Potencial químico

La búsqueda restringida de Levy (Parr y Yang, 1989) permite expresar la energía como un funcional de la densidad en base a los teoremas de Hohenberg-Kohn, esto es:

$$E[\rho] = F[\rho] + \int v(\boldsymbol{r})\rho\,(\boldsymbol{r}) d\boldsymbol{r}$$

Así mismo, a temperatura finita se encuentran tanto la energía libre de Helmholtz a partir de la teoría del ensamble-canónico

$$A[\rho] = F[\rho] + \int v(\boldsymbol{r})\rho\,(\boldsymbol{r}) d\boldsymbol{r}$$

como el gran potencial a partir de la teoría del ensamble-gran-canónico

$$\Omega[\rho(\boldsymbol{r})] = F[\rho(\boldsymbol{r})] + \int \rho\,(\boldsymbol{r})(v(\boldsymbol{r}) - \mu) d\boldsymbol{r}$$

Dada la $v - N -$ representabilidad de la densidad electrónica y la búsqueda restringida (Parr y Yang, 1989), se define

$$F[\rho(\boldsymbol{r})] = \underset{\hat{\Gamma}_N \to \rho}{\text{Min}}\, tr\left[\hat{\Gamma}_N\left(\hat{T}[\rho] + \hat{U}[\rho] - \frac{1}{\beta} ln\hat{\Gamma}_N\right)\right]$$

Por medio del principio variacional el funcional $F[\rho]$ se expresa como

$$F[\rho] = T[\rho] + U[\rho] - \theta S[\rho]$$

el funcional de energía cinética

$$T[\rho] = tr\left(\hat{\Gamma}_{N,\rho}^{min}\hat{T}\right)$$

el funcional de energía potencial de interacción partícula-partícula

$$U[\rho] = tr\left(\hat{\Gamma}_{N,\rho}^{min}\hat{U}\right)$$

y el funcional de entropía

$$S[\rho] = -k_B tr\left(\hat{\Gamma}_{N,\rho}^{min} ln\hat{\Gamma}_{N,\rho}^{min}\right)$$

donde $\hat{\Gamma}_{N,\rho}^{min}$ es el operador de densidad, k_B es la constante de Boltzmann y θ es la temperatura absoluta.

Aquí el interés se enfoca en el límite a temperatura-cero, para que en $\theta = 0$ $(\beta = \infty)$, el estado de equilibrio y el estado base sean el mismo, esto es, el estado principal. Conforme $\theta \to 0$, el término $\theta S[\rho]$ desaparece de $A[\rho]$ y de $\Omega[\rho]$.

En el límite cuando $\beta \to 0$ el funcional universal $F[\rho]$ se convierte en

$$F_{GC}[\rho(\boldsymbol{r})] = \underset{\Gamma\to\rho}{\text{Min}}\, Tr\left[\hat{\Gamma}\left(\hat{T} + \hat{V}_{ee}\right)\right]$$

Donde F_{GC} es el funcional universal del estado-base, en el cual se extiende la búsqueda restringida de Levy a operadores de densidad general $\hat{\Gamma}$ en el espacio de Fock (ensamble gran canónico). El potencial $\hat{V}_{ee}$ se escribe en lugar del potencial de interacción partícula-partícula $\hat{U}$ debido a que la esencia está en los potenciales electrónicos. El gran potencial en $\beta \to \infty$ correspondiente, se reduce a

$$\Omega[\rho(\boldsymbol{r})] = F_{GC}[\rho] + \int \rho\,(\boldsymbol{r})(v(\boldsymbol{r}) - \mu)d\boldsymbol{r}$$

$$= E_{GC}[\rho] - \mu N$$

En el cual entra la energía funcional del ensamble gran-canónico

$$E_{GC}[\rho] = F_{GC}[\rho] + \int \rho\,(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r}$$

Tanto $\Omega[\rho]$ como $E_{GC}[\rho]$ están definidos para densidades que integran a cualquier número positivo finito.

En la construcción del gran potencial por búsqueda restringida, del principio variacional cuando $\beta \to \infty$ se tiene que

$$\Omega^0 = \mathrm{E}^0(N) - \mu N \leq Tr\hat{\Gamma}\left(\widehat{H} - \mu\widehat{N}\right)$$

Donde $\mathrm{E}^0(N)$ y μ son la energía y el potencial químico del estado base con N electrones. Entonces, el principio variacional del funcional-de-la-densidad correspondiente es

$$\mathrm{E}^0(N) - \mu N \leq E_{GC}[\rho(\boldsymbol{r})] - \mu \int \rho\,(\boldsymbol{r})d\boldsymbol{r}$$

De la cual se sigue la ecuación variacional para la densidad y la energía del estado-base del electrón

$$\delta\left\{E_{GC}[\rho(\boldsymbol{r})] - \mu \int \rho\,(\boldsymbol{r})d\boldsymbol{r}\right\} = 0$$

entonces la ecuación de Euler-Lagrange es

$$\frac{\delta E_{GC}[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})} - \mu = 0$$

Equivalentemente

$$\frac{\delta F_{GC}[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})} + v(\boldsymbol{r}) - \mu = 0$$

En la última ecuación se ha supuesto que $F_{GC}[\rho(\boldsymbol{r})]$ es diferenciable. Ahora se muestra que Ω^0 no es verdadero en general; sin embargo, es válido para sistemas atómicos y moleculares. Consideremos una $\hat{\Gamma}$ de prueba que describe un estado base con el número promedio de electrones $N + \Delta N = Tr\hat{\Gamma}\widehat{N}$. Entonces se tiene que

$$\mathrm{E}^0(N) - \mu N \leq \mathrm{E}^0(N + \Delta N) - \mu(N + \Delta N)$$

Análogamente, cuando la $\hat{\Gamma}$ de prueba es $Tr\hat{\Gamma}\widehat{N} = N - \Delta N$

$$\mathrm{E}^0(N) - \mu N \leq \mathrm{E}^0(N - \Delta N) - \mu(N - \Delta N)$$

Sumando estas dos expresiones se tiene

$$2\mathrm{E}^0(N) - 2\mu N \leq \mathrm{E}^0(N + \Delta N) + \mathrm{E}^0(N - \Delta N) - \mu(N + \Delta N) - \mu(N - \Delta N)$$

$$\therefore \qquad \mathrm{E}^0(N + \Delta N) + \mathrm{E}^0(N - \Delta N) - 2\mathrm{E}^0(N) \geq 0$$

Que demanda que la función $\mathrm{E}^0(N)$ sea convexa. En consecuencia, si $E^0(N)$ es una función diferenciable

$$\mu = \frac{\partial E^0(N)}{\partial N}$$

y si $E^0(N)$ es una función dos veces diferenciable, entonces

$$\frac{\partial^2 E^0(N)}{\partial N^2} \geq 0$$

Sin embargo, poniendo $\Delta N = 1$ se tiene que

$$\mathrm{E}^0(N+1) + \mathrm{E}^0(N-1) - 2\mathrm{E}^0(N) \geq 0$$

$$\mathrm{E}^0(N+1) - \mathrm{E}^0(N) \geq \mathrm{E}^0(N) - \mathrm{E}^0(N-1)$$

O bien

$$I(N+1) \leq I(N)$$

Donde $I(N)$ es el potencial de ionización del estado base de N-electrones. Esta ecuación dice que los potenciales de ionización sucesivos no son decrecientes (para potencial externo fijo). Una tendencia que se ha observado para el caso de iones de oxígeno es que la *energía se comporta como una función del número de electrones* (Phillips y Davidson, 1984).

$$E(N) = -\sum_{M=1}^{N} I(M)$$

Por lo tanto, el potencial químico en estas ecuaciones es el límite de temperatura-cero del potencial químico definido para el ensamble del gran-potencial a temperatura finita,

$$\mu = \left(\frac{\partial A}{\partial N}\right)_{\theta, v(r)} = \left(\frac{\partial E}{\partial N}\right)_{\theta, v(r)} - \theta \left(\frac{\partial S}{\partial N}\right)_{\theta, v(r)}$$

Donde A es la energía libre de Helmholtz. El símbolo μ se usa para designar el límite a temperatura-cero:

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{v} = \lim_{\theta \to 0} \left(\frac{\partial A}{\partial N}\right)_{\theta, v(r)}$$

Y se simplifica la notación al tomar

$$\delta\{E[\rho] - \mu N(\rho)\} = 0$$

y

$$\mu = \left[\frac{\delta E}{\delta \rho}\right]_{v} = v(\boldsymbol{r}) + \frac{\delta F[\rho]}{\delta \rho(\boldsymbol{r})}$$

Estas son las ecuaciones básicas del estado base en la teoría de funcionales de la densidad para átomos, moléculas y sólidos.

El potencial químico de la teoría funcional-de-la-densidad mide la tendencia de los electrones a escapar de una nube electrónica. Es una constante en todo el espacio, para el estado base de un átomo, molécula o sólido, e igual a la pendiente de la curva E vs N a $v(\boldsymbol{r})$ constante. El potencial químico es el negativo de la electronegatividad (Parr et al., 1978). Para la aproximación de diferencias-finitas de tres puntos a $\frac{\partial E}{\partial N}$.

Para una especie S es

$$\mu = -\frac{I + A}{2}$$

Donde $I = E_S^+ - E_S \;\; y \;\; A = E_S - E_S^-$ son respectivamente el potencial de ionización y la afinidad electrónica para las especies. En acuerdo con el teorema de Koopmans

$$I = E(N-1) - E(N)$$

$$A = E(N) - E(N+1)$$

La fórmula de Mulliken para electronegatividad es

$$x_M = \frac{I + A}{2}$$

$$\therefore \qquad \mu \approx -x_M$$

y el concepto de potencial químico es el mismo que el concepto de electronegatividad. Adicionalmente, el potencial químico también se identifica en la teoría del ensamble canónico. La manera de hacer esto es a través del método de diferencias finitas. Para el sistema de N electrones, se toma

$$\mu^- = \frac{E(N-1) - E(N)}{-1} = -I \quad (pendiente\ cuando\ el\ sistema\ saca\ un\ electrón)$$

$$\mu^+ = E(N+1) - E(N) = -A \quad (pendiente\ cuando\ el\ sistema\ agrega\ un\ electrón)$$

$$\mu^0 = -\frac{I + A}{2} \quad (pendiente\ cuando\ un\ electrón\ ni\ se\ saca\ ni\ se\ agrega)$$

También, la curvatura de $E(N)$ se convierte en

$$\eta = \frac{E(N+1) - E(N) - \big(E(N) - E(N-1)\big)}{2} = \frac{I - A}{2}$$

Cuando $\beta \to \infty$ entonces la temperatura tiende a cero, y solo los estados base de algunas especies sobreviven: i) el estado base neutral con energía E_0 y número de electrones N_0, ii)

un estado base ión-positivo con energía $E_0 + I$ y número de electrones $N_0 - 1$, y iii) un estado base ion-negativo con energía $E_0 - A$ y número de electrones $N_0 + 1$. Todos los estados se suponen no-degenerados. Consideremos

$$E(N_0) = E_{N_0} = E_0$$

de acuerdo con el potencial gran-canónico, si $N = N_0$ entonces

$$\mu_0 = \frac{E^0_{N_0-1} - E^0_{N_0+1}}{2} + degeneración$$

Si $N = N_0 + \Delta N$ donde N_0 es algún entero y ΔN es positivo, entonces

$$\mu_0 = E^0_{N_0} - E^0_{N_0+1}$$

para $N_0 + 1 > N > N_0$. Y si ΔN es negativo, entonces

$$\mu_0 = E^0_{N_0} - E^0_{N_0-1}$$

para $N_0 - 1 < N < N_0$. Es decir

$$\mu_0 = \begin{cases} -I & N_0 - 1 < N < N_0 \\ -\dfrac{I + A}{2} & N = N_0 \\ -A & N_0 + 1 > N > N_0 \end{cases}$$

La energía de ionización y la afinidad electrónica se pueden aproximar como las energías HOMO y LUMO, respectivamente. Nótese que si sumamos las contribuciones para ΔN positivo, y ΔN negativo se tiene que

$$2\mu_0 = 2E^0_{N_0} - \left(E^0_{N_0+1} + E^0_{N_0-1}\right)$$

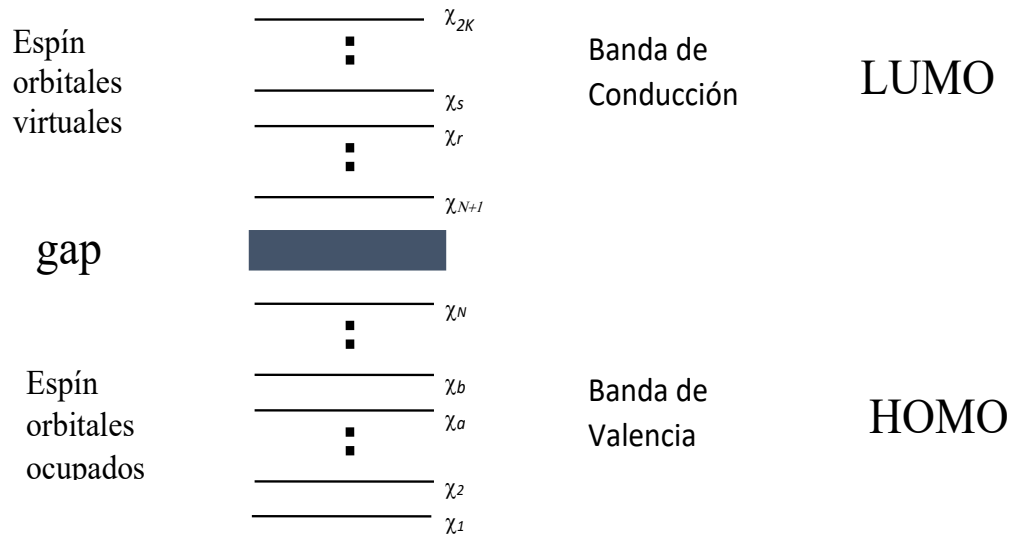
$$\mu_0 = E^0_{N_0} - \frac{E^0_{N_0+1} + E^0_{N_0-1}}{2}$$

$$\therefore \qquad \mu = \mu_0 - E^0_{N_0} = -\frac{E_{LUMO} + E_{HOMO}}{2}$$

Esto se ilustra con el esquema de la figura 36, donde

$$E_{gap} = E_{LUMO} - E_{HOMO}$$

Figura 36. *Ilustración de las bandas*



Fuente: elaboración propia.

La curvatura de la energía *E* como una función del número de electrones *N* a 0 K es otra cantidad de interés, que se obtiene derivando dos veces la función de la energía con respecto al número de electrones

$$2\eta = \left(\frac{\partial^2 E}{\partial N^2}\right)_{\theta \to 0} = \lim_{\beta \to \infty} \left[\frac{\frac{\partial}{\partial \mu}\left(\frac{\partial E}{\partial N}\right)_{\beta}}{\left(\frac{\partial N}{\partial \mu}\right)_{\beta}}\right]$$

Se encuentra que

$$2\eta = \begin{cases} 0 & N < N_0 \\ \lim_{\beta \to \infty}\left[\eta_0 e^{\beta \eta_0}\right] = \infty & N = N_0 \\ 0 & N > N_0 \end{cases}$$

donde

$$\eta_0 = \frac{I - A}{2}$$

Es la dureza absoluta.

Ahora bien, Geerlings et al. (2003) definen el multiplicador lagrangiano μ como el negativo de la electronegatividad

$$\mu = -x = -\left(\frac{\partial E}{\partial N}\right)_v$$

y la electronegatividad se calcula como el promedio de las derivadas del lado izquierdo y del lado derecho, esto es

$$x^- = E(N = N_0 - 1) - E(N = N_0) = I$$

$$x^+ = E(N = N_0) - E(N = N_0 + 1) = A$$

$$x = \frac{1}{2}(x^+ + x^-) = \frac{1}{2}(I + A)$$

donde I y A son la energía de ionización y la afinidad electrónica del sistema estudiado $N_0 - electrón$ (neutro o cargado), y aprovechando el teorema de Koopmans,

$$x = \frac{1}{2}(\epsilon_{HOMO} + \epsilon_{LUMO})$$

Cuando se evalúa la electronegatividad entre $q = -e \; y \; q = +e$ se obtiene la electronegatividad de Mulliken $x = (I + A)/2$, al igual que si se evalúa la dureza análogamente se obtiene

$$\eta = \frac{I - A}{2}$$

$$\eta = \frac{1}{2}(\epsilon_{LUMO} - \epsilon_{HOMO})$$

Métodos del gradiente y propiedades moleculares

La geometría de equilibrio de una molécula corresponde al arreglo nuclear que minimiza la energía. Los cambios en la energía como una longitud o ángulo de enlace que varían en intervalos moderados son notablemente menores que la energía de correlación molecular. A pesar de la pequeñez de estos cambios la energía molecular con localizaciones nucleares comparada con el error en la energía (la energía de correlación) inherente en el método SCF, la función de onda *ab initio* usualmente da buenas predicciones (0 a 3% de error) de los ángulos y distancias de enlace de equilibrio en moléculas que no involucran metales de transición. Evidentemente, la energía de correlación permanece aproximadamente constante para variaciones de ángulos y longitudes de enlace en la región de geometría de equilibrio.
Para una molécula diatómica, la energía potencial molecular *E* depende solo de la distancia internuclear *r*. Entonces, para encontrar el potencial mínimo se necesita localizar un cero en

dE/dr. La búsqueda es más complicada para moléculas diatómicas debido a que la energía potencial es una función de muchas coordenadas nucleares q_i. En la geometría de equilibrio, cada una de las fuerzas f_i ejercida sobre un núcleo por electrones y otros núcleos debe desaparecer:

$$f_i = -\frac{\partial E}{\partial q_i} = 0$$

Por lo tanto, en principio, la geometría de equilibrio se puede encontrar computando todas las fuerzas en una geometría molecular dada y viendo si desaparecen. Si no desaparecen, variamos la geometría hasta que corresponda a fuerzas cero (un gradiente vectorial de longitud cero). Computacionalmente, las fuerzas no desaparecen idénticamente, pero calculando sus magnitudes podemos parar la búsqueda iterativa para la geometría de equilibrio cuando las magnitudes son cercanas a cero. Los valores óptimos de equilibrio de ángulos y distancias de enlace computados en 1° y $0.002a_0$, deberían estar debajo de 10 pN para todas las fuerzas.

Un gradiente cero caracteriza un punto estacionario en la superficie, pero no consigue diferenciar entre mínimo, máximo o punto silla. Por lo tanto, el procedimiento de búsqueda nos identifica no solo una geometría de equilibrio sino también el estado transicional de la reacción química, esta última corresponde a un punto silla en la superficie. Para distinguir los tipos de puntos estacionarios es necesario considerar las segundas derivadas de la energía con respecto a las coordenadas nucleares. Las cantidades $\partial^2 E/\partial q_i \partial q_j$ comprenden la matriz hessiana. Mientras que un mínimo (máximo) de una curva corresponde a una segunda derivada positiva (negativa), un mínimo (máximo) de una superficie de energía potencial multidimensional se caracteriza porque todos los eigenvalores de la matriz hessiana sean positivos (negativos). Un estado de transición (un punto-silla de primer orden) corresponde a un eigenvalor negativo y todos los demás positivos.

Optimización de geometría

Existen muchos procedimientos matemáticos para encontrar un mínimo local o una función de varias variables. Estos procedimientos encuentran un mínimo local de la energía en la vecindad de la geometría inicialmente supuesta. Al proceso de encontrar tal mínimo se le llama *optimización de geometría* o *minimización de la energía*. Para una molécula con varias

conformaciones, uno debe repetir el procedimiento de búsqueda del mínimo-local para cada conformación posible, hasta localizar el mínimo global.

Los procedimientos eficientes para encontrar un mínimo local de la energía requieren cálculos repetidos tanto de la energía como de sus derivadas. El conjunto de 3N-6 primeras derivadas parciales de la energía con respecto a cada una de sus variables constituye un vector (en un espacio de 3N-6 dimensiones) llamado el gradiente de la energía. En un mínimo local, que el gradiente sea cero significa que cada una de las primeras 3N-6 derivadas parciales de la energía debe ser cero. Cualquier punto en la superficie de energía potencial (PES) donde el gradiente es cero y, es llamado un punto estacionario o punto crítico. Un punto estacionario en una PES puede ser un mínimo, un máximo, o un punto silla.

Cuando se arregla el conjunto $\partial^2 E/\partial q_1^2$, $\partial^2 E/\partial q_1 \partial q_2$, $\partial^2 E/\partial q_2 \partial q_1$, $\partial^2 E/\partial q_2^2$,…, de segundas derivadas en una matriz cuadrada conocida como *hessiano* o *matriz de fuerza constante*, una manera eficiente de encontrar el mínimo local de una función de varias variables es el método de *Newton* (o *Newton-Raphson*) que aproxima la función por una expansión en serie de Taylor que termina en los términos cuadráticos y usa las primeras y segundas derivadas parciales de la función exactamente evaluadas.

Algoritmos de optimización. Los principios básicos de optimización fueron redescubiertos para implementarse en computadoras digitales. Algunos de estos algoritmos se bosquejan a continuación.

Sistemas de gradiente dinámico

a) Algoritmo básico iterativo de un gradiente descendente

El método descendente, también conocido como método del gradiente, es una de las técnicas más antiguas para minimizar una función determinada definida en un espacio n dimensional. Este método es la base para métodos directos usados en optimización de problemas restringidos y no restringidos. Es la técnica más frecuentemente usada para la optimización no lineal. En otras palabras, lo que se busca es un valor de x que minimice la función $E(\boldsymbol{x})$. Donde $E(\boldsymbol{x}) = E(x_1, x_2, \dots, x_n)$ es una función de valor real de n variables a ser minimizada (o maximizada) para lo cual se tienen que observar sus valores en los puntos críticos, esto es, en puntos $\boldsymbol{x}\epsilon$R en los cuales el problema de optimización a considerar es

$$\nabla E(\boldsymbol{x}) = \mathbf{0}$$

y suponiendo que el gradiente ∇E del vector E y la *matriz Hessiana* de la función objetivo existen; es decir, pueden ser evaluados analíticamente, entonces el método iterativo implica generar una secuencia de puntos de búsqueda $x^{(k)}$ a través del procedimiento iterativo

$$x^{k+1} = x^{(k)} + \eta^{(k)} d_k, \qquad x(0) = x^{(0)} \qquad (k = 0,1,2,\dots,)$$

Donde el escalar $\eta^{(k)} > 0$ determina la longitud del paso (razón de aprendizaje) a ser tomado en la dirección del vector $d_k \cong \Delta x^{(k)}$. En la optimización numérica existen distintas técnicas para el cálculo del parámetro $\eta^{(k)}$ y la dirección del vector d_k que son conocidos. Además del método del gradiente descendiente se consideran otros tres métodos básicos:

1. Para el método del gradiente (*steepest-descent*) definimos la dirección como:

$$d_k = \nabla E\left(x^{(k)}\right)$$

2. Para el método de Newton la dirección de búsqueda está determinada por la ecuación:

$$d_k = \left[\nabla^2 E\left(x^{(k)}\right)\right]^{(-1)} \nabla E\left(x^{(k)}\right)$$

3. Ya que el cálculo de la *matriz inversa hessiana* $\left[\nabla^2 E\left(x^{(k)}\right)\right]^{(-1)}$ puede ser ligeramente complejo, se aproxima a una matriz positiva definida simétrica de dimensión $n \times n$, llamada H_k. Aplicando el método de cuasi-Newton (variable métrica) la dirección de búsqueda se determina por:

$$H_k = -\left[\nabla^2 E\left(x^{(k)}\right)\right]^{(-1)}$$

$$d_k = -H_k \nabla E\left(x^{(k)}\right)$$

4. El método del gradiente conjugado calcula la dirección de búsqueda actual d_k como una combinación lineal del gradiente del vector actual y la anterior búsqueda de la dirección. Esta forma simple de buscar la dirección se calcula por:

$$d_k = -\nabla E\left(x^{(k)}\right) + \beta_k d_{k+1} \quad (k = 0,1,2,\dots,)$$

$$con\ \ d_0 = -\nabla E\left(x^{(0)}\right)$$

donde β_k es un parámetro escalar que asegura que la secuencia de vectores d_k satisfaga la condición de conjugación mutua.

La longitud del paso $\eta^{(k)}$ generalmente se determina en estos métodos usando dos técnicas: *i*) minimización a lo largo de la línea y *ii*) usando un valor fijo de la longitud de paso en una dimensión.

b) Método del gradiente descendente

El método del gradiente descendente se puede escribir de forma general como:

$$x^{(k+1)} = x^{(k)} - M_k \nabla E\left(x^{(k)}\right),$$

donde M_k es una matriz simétrica, positiva definida, de dimensión $n \times n$. La elección apropiada de la matriz M_k es crítica en vista de las propiedades de convergencia del algoritmo. El algoritmo de minimización determina el mínimo local de la función $E(x)$ como el límite de la secuencia $E\left(x^{(k)}\right)$ (k=0,1,2, …) donde $x^{(0)}$ es una estimación inicial del minimizador local x^*.

Estos algoritmos iterativos generan una secuencia de puntos $\{x(k)\}$ y una dirección de búsqueda $d_k \cong \Delta x^{(k)}$, lo cual puede ser a través de una aproximación definida discreta para alguna trayectoria de tiempo continuo desde el punto de inicio $x^{(0)}$ al punto mínimo x^* (estacionario). La trayectoria de tiempo continuo $x^{(t)}$ se determina por lo general con un sistema de ecuaciones diferenciales de la forma:

$$\frac{dx_j}{dt} = -\sum_{i=1}^{n} \mu_{ji} \frac{\partial E(x)}{\partial x_i}, \qquad x_i(0) = x_i^{(0)} \qquad (j = 0,1,2,3, \dots, n)$$

La cual puede ser escrita de manera más compacta en forma de matriz

$$dx/dt = -\mu \nabla E(x), \qquad x(0) = x^{(0)}$$

donde $x(t) \epsilon \mathbb{R}^n \ y \ \mu(x,t)$ es una matriz positiva definida de dimensión $n \times n$ cuyas entradas son generalmente dependientes del tiempo y la variable $x(t) = \left(x_1(t), x_2(t), \dots, x_n(t)\right)$. La definición de la matriz positiva *m* se requiere para asegurar la estabilidad del sistema.

c) Método de Newton

Para construir el algoritmo que nos lleva al método de un punto fijo apropiado en el caso unidimensional, se obtiene la función ϕ con la propiedad de que

$$g(x) = x - \phi(x) f(x)$$

da una convergencia cuadrática en el punto fijo p de la función g. A partir de esta condición el método de Newton evoluciona al seleccionar $\phi(x) = {}^{1}/_{f'(x)}$, suponiendo que $f'(x) \neq 0$

La aplicación de un procedimiento semejante en el caso *n*-dimensional incluye una matriz

$$A(\boldsymbol{x}) = \begin{bmatrix} a_{11}(\boldsymbol{x}) & a_{12}(\boldsymbol{x}) & \cdots & a_{1n}(\boldsymbol{x}) \\ a_{21}(\boldsymbol{x}) & a_{22}(\boldsymbol{x}) & \cdots & a_{2n}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \ \vdots \\ a_{n1}(\boldsymbol{x}) & a_{n2}(\boldsymbol{x}) & \cdots & a_{nn}(\boldsymbol{x}) \end{bmatrix} \qquad (92)$$

donde todos los elementos $a_{ij}(\boldsymbol{x})$ son una función de $\mathbb{R}^n$ en $\mathbb{R}$. Esto requiere obtener A(**x**) de modo que

$$G(\boldsymbol{x}) = \boldsymbol{x} - A(\boldsymbol{x})^{-1}\boldsymbol{F}(\boldsymbol{x})$$

da la convergencia cuadrática a la solución de $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}$, suponiendo que A($\boldsymbol{x}$) es no singular en el punto fijo **p** de **G**.

Teorema. Supongamos que **p** es una solución de **G**(**x**)=**x** para alguna función $\boldsymbol{G}(\boldsymbol{x}) = (x_1, x_2, \dots, x_n)$ que mapea $\mathbf{R}^n$ en $\mathbf{R}^n$. Si existe un número δ>0 con la propiedad de que

I. $\frac{\partial g_i}{\partial x_j}$ sea continua en $N_\delta = \{\boldsymbol{x} | \|\boldsymbol{x} - \boldsymbol{p}\| < \delta\}$ para toda $i = 1,2,\dots,n \;\; y \;\; j = 1,2,\dots,n$;

II. $\frac{\partial^2 g_i(\boldsymbol{x})}{\partial x_j \partial x_k}$ sea continua y $\left|\frac{\partial^2 g_i(\boldsymbol{x})}{\partial x_j \partial x_k}\right| \leq M$ para alguna constante M siempre que $x \epsilon N_\delta$ para toda $i = 1,2,\dots,n\,, \quad j = 1,2,\dots,n \quad y \quad k = 1,2,\dots,n$;

III. $\frac{\partial g_i(\boldsymbol{p})}{\partial x_k} = 0$ para toda $i = 1,2,\dots,n \;\; y \;\; k = 1,2,\dots,n$;

Entonces existe un número $\hat{\delta} \leq \delta$ tal que la sucesión generada por $\boldsymbol{x}^k = G\left(\boldsymbol{x}^{(k-1)}\right)$ converge cuadráticamente a **p** para cualquier elección de $\boldsymbol{x}^{(0)}$ a condición de que $\left\|\boldsymbol{x}^{(0)} - \boldsymbol{p}\right\| < \delta$. Más aun,

$$\left\|\boldsymbol{x}^{(0)} - \boldsymbol{p}\right\|_\infty \leq \frac{n^2 M}{2}\left|\left|\boldsymbol{x}^{(k-1)} - \boldsymbol{p}\right|\right|_\infty^2, \qquad para\ toda\ \ k \geq 1$$

Para utilizar el teorema anterior, suponga que $A(\boldsymbol{x})$ es una matriz de $n \times n$ de funciones de $\mathbb{R}^n$ a $\mathbb{R}$ en la forma de la ecuación (92), cuyos elementos específicos se escogerán adelante. Suponga además que $A(\boldsymbol{x})$ es no singular cerca de una solución $\boldsymbol{p}$ de $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}$, y denotemos con $b_{ij}(\boldsymbol{x})$ el elemento de $A(\boldsymbol{x})^{-1}$ en el *i*-ésimo renglón y en la *j*-ésima columna.

Dado que $\boldsymbol{G}(\boldsymbol{x}) = \boldsymbol{x} - \mathrm{A}(\boldsymbol{x})^{-1}\mathbf{F}(\boldsymbol{x})$, tenemos $g_i(\boldsymbol{x}) = x_i - \sum_{j=1}^{n} b_{ij}(\boldsymbol{x})\, f_j(\boldsymbol{x})$ y

$$\frac{\partial g_i(\boldsymbol{x})}{\partial x_k} = \begin{cases} 1 - \sum_{j=1}^{n} \left(b_{ij}(\boldsymbol{x}) \frac{\partial f_j}{\partial x_k}(\boldsymbol{x}) + \frac{\partial b_{ij}}{\partial x_k}(\boldsymbol{x}) f_j(\boldsymbol{x}) \right), & si \quad i = k \\ -\sum_{j=1}^{n} \left(b_{ij}(\boldsymbol{x}) \frac{\partial f_j}{\partial x_k}(\boldsymbol{x}) + \frac{\partial b_{ij}}{\partial x_k}(\boldsymbol{x}) f_j(\boldsymbol{x}) \right), & si \quad i \neq k \end{cases}$$

El teorema implica que se necesita $\frac{\partial g_i}{\partial x_k}(\boldsymbol{p})$ para toda $i = 1,2,\dots,n \;\; y \;\; k = 1,2,\dots,n$. Esto significa que, cuando $i = k$,

$$1 - \sum_{j=1}^{n} b_{ij}(\boldsymbol{p}) \frac{\partial f_j}{\partial x_k}(\boldsymbol{p}) = 0$$

$$\therefore \qquad \sum_{j=1}^{n} b_{ij}(\boldsymbol{p}) \frac{\partial f_j}{\partial x_i}(\boldsymbol{p}) = 1$$

cuando $k \neq i$,

$$-\sum_{j=1}^{n} b_{ij}(\boldsymbol{p}) \frac{\partial f_j}{\partial x_i}(\boldsymbol{p}) = 0$$

$$\therefore \qquad \sum_{j=1}^{n} b_{ij}(\boldsymbol{p}) \frac{\partial f_j}{\partial x_i}(\boldsymbol{p}) = 0$$

Al definir la matriz $J(\boldsymbol{x})$ por medio de

$$J(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x}) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\boldsymbol{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_n}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}) \end{bmatrix}$$

para la cual se requiere la matriz identidad $\mathrm{A}(\boldsymbol{p})^{-1} J(\boldsymbol{p}) = I$,

$$\therefore \qquad \mathrm{A}(\boldsymbol{p}) = J(\boldsymbol{p})$$

En consecuencia, una elección apropiada de $A(\boldsymbol{x})$ es $\mathrm{A}(\boldsymbol{x}) = J(\boldsymbol{x})$ dado que entonces se cumple la condición (III) del teorema.

La función **G** está definida por $\boldsymbol{G}(\boldsymbol{x}) = \boldsymbol{x} - J(\boldsymbol{x})^{-1}\boldsymbol{F}(\boldsymbol{x})$ y el procedimiento de la iteración funcional pasa de seleccionar $x^{(0)}\ para\ k \geq 1$, a generar

$$\boldsymbol{x}^{(k)} = \boldsymbol{G}\left(\boldsymbol{x}^{(k-1)}\right) = \boldsymbol{x}^{(k-1)} - J\left(\boldsymbol{x}^{(k-1)}\right)^{-1}\boldsymbol{F}\left(\boldsymbol{x}^{(k-1)}\right)$$

A esto se le llama método de Newton para sistemas no lineales y generalmente se espera que dé una convergencia cuadrática, siempre y cuando se conozca un valor inicial suficientemente preciso y exista $J(\boldsymbol{p})^{-1}$.

d) Método cuasi-Newton

Un punto débil importante del Método de Newton en que la solución de los sistemas de ecuaciones no lineales es el requisito de que en cada iteración es necesario calcular una matriz Jacobiana, y resolver un sistema lineal de $n \times n$ que la contiene. Para ejemplificar la importancia de tal debilidad, consideremos los cálculos que conlleva una iteración de dicho método. La matriz jacobiana asociada a un sistema de *n* ecuaciones no lineales de la forma $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}$, requiere determinar y evaluar las n^2 derivadas parciales de las *n* funciones componentes de $\boldsymbol{F}$.

Cuando no es práctico efectuar la evaluación exacta, podemos usar las aproximaciones de diferencia finita a las derivadas parciales. Por ejemplo

$$\frac{\partial f_j}{\partial x_k}\left(\boldsymbol{x}^{(i)}\right) \approx \frac{f_j\left(\boldsymbol{x}^{(i)} + \boldsymbol{e}_k h\right) - f_j\left(\boldsymbol{x}^{(i)}\right)}{h}$$

donde *h* es pequeña en valor absoluto y $\boldsymbol{e}_k$ es el vector cuyo único elemento distinto de cero es un 1 de la *k*-ésima coordenada. El número total de cálculos que se requieren para una sola evaluación del método de Newton es $n^2 + n$ evaluaciones de funciones escalares por iteración. Pertenece a una clase de técnicas denominadas actualizaciones de secante (método de Broyden) con cambio mínimo que dan origen a los algoritmos **cuasi-Newton**. Estos métodos reemplazan la matriz jacobiana con una matriz de aproximación que se actualiza en

cada iteración. Su desventaja radica en que se pierde la convergencia cuadrática de Newton, al ser sustituida por una convergencia denominada **superlineal**, la cual implica que

$$\lim_{i\to\infty} \frac{\|\boldsymbol{x}^{(i+1)} - \boldsymbol{p}\|}{\|\boldsymbol{x}^{(i)} - \boldsymbol{p}\|} = 0$$

donde $\boldsymbol{p}$ denota la solución de $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}$ y $\boldsymbol{x}^{(i)}$ y $\boldsymbol{x}^{(i+1)}$ son aproximaciones consecutivas de $\boldsymbol{p}$. El método procede de manera semejante al método de Newton, porque se reemplaza la matriz $J(\boldsymbol{x}^{(1)})$ por una matriz A_1 que tiene la propiedad

$$A_1(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}) = \boldsymbol{F}(\boldsymbol{x}^{(1)}) - \boldsymbol{F}(\boldsymbol{x}^{(0)})$$

Todo vector distinto de cero en $\mathbb{R}^n$ puede escribirse como la suma de un múltiplo de $\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}$ y de un múltiplo de un vector en el complemento ortogonal de $\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}$. Entonces, si queremos definir a la matriz A_1, tenemos que determinar cómo actúa sobre el complemento ortogonal de $\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}$. Esto es

$$A_1(\boldsymbol{z}) = J(\boldsymbol{x}^{(0)})\boldsymbol{z}, \text{siempre que } (\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)})^t \boldsymbol{z} = 0.$$

Entonces, la única manera de definir A_1 es

$$A_1 = J(\boldsymbol{x}^{(0)}) + \frac{[\boldsymbol{F}(\boldsymbol{x}^{(1)}) - \boldsymbol{F}(\boldsymbol{x}^{(0)}) - J(\boldsymbol{x}^{(0)})(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)})](\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)})^t}{\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\|_2^2}$$

Esta matriz es la que se usa en lugar de $J(\boldsymbol{x}^{(1)})$ para determinar $\boldsymbol{x}^{(2)}$ como

$$\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} - A_1^{-1}\boldsymbol{F}(\boldsymbol{x}^{(1)})$$

En general, una vez determinado $\boldsymbol{x}^{(i)}$, se calcula $\boldsymbol{x}^{(i+1)}$ usando las siguientes expresiones

$$A_i = A_{i+1} + \frac{\boldsymbol{y_i} - A_{i-1}\boldsymbol{s}_i}{\|\boldsymbol{s}_i\|_2^2}$$

$$y \qquad \boldsymbol{x}^{(i+1)} = \boldsymbol{x}^{(i)} - A_1^{-1}\boldsymbol{F}(\boldsymbol{x}^{(i)})$$

donde la notación $\boldsymbol{y_i} = \boldsymbol{F}(\boldsymbol{x}^{(i)}) - \boldsymbol{F}(\boldsymbol{x}^{(i+1)})$ y $\boldsymbol{s}_i = \boldsymbol{x}^{(i)} - \boldsymbol{x}^{(i+1)}$ es para simplificar solamente.

e) Gradiente conjugado

Es un gran método para matrices simétricas y positivas definidas. Ya que aprovecha muy bien la estructura de la matriz y tiene muy buenas propiedades de estabilidad numérica.

El método del gradiente conjugado es un método iterativo en que a partir de un iterante inicial va calculando iterantes sucesivos que se van acercando a la solución exacta del sistema lineal. El iterante k+1 será la solución si la diferencia entre él y el iterante k es menor que un cierto número prefijado. El método del gradiente conjugado se enfoca como un método de minimización del funcional J, convexo y con un único mínimo:

$$J(u) = \frac{1}{2}(Au, u) - (b, u), \qquad \forall u \epsilon \mathbb{R}^n$$

siendo (x, y) el producto interior de x por y.

Así para k = (1, ..., N-1), se calcula el mínimo del funcional J sobre la variedad lineal $x^0 + (d^0, \ldots, d^N)$

Esto es, sea $x^0 \varepsilon \mathbb{R}^N$, constrúyase la base $(d^0, d^1, \ldots, d^N)$ ortogonal de $\mathbb{R}^N$ respecto al producto escalar (A ▪ , ▪). Al realizar una serie de cuentas, el algoritmo nos queda de la siguiente forma:

a) Inicialización del algoritmo: sea el iterante inicial $x^0 \varepsilon \mathbb{R}^N$ cualquiera, definimos $r^0 = b - Ax^0$ y definimos la primera dirección de descenso como $d^0 = r^0$.

b) Iteraciones: para k = 0, . . . , N − 1 se hace la minimización correspondiente, esta equivale a: sea $a^k = \frac{(r^k, r^k)}{Ad^k, d^k}$, sea $x^{k+1} = x^k + a^k d^k$ el siguiente iterante y sea $r^{k+1} = r^k - a^k A d^k$ el siguiente residuo. Posteriormente, las iteraciones se completan calculando la siguiente dirección de descenso:

$$\beta^{k+1} = \frac{(r^{k+1}, r^{k+1})}{(r^k, r^k)}$$

y la siguiente dirección de descenso seria, $d^{k+1} = r^{k+1} + \beta^{k+1} d^k$.

Se puede probar que las direcciones de descenso así definidas son conjugadas, es decir, ortogonales 2 a 2 con el producto escalar (A ▪ , ▪) y que este algoritmo nos va a llevar al mínimo de *J* y por tanto a la solución del problema. Además, los residuos son ortogonales entre sí.

El algoritmo del *gradiente conjugado*, así definido, converge a lo sumo en N iteraciones a la solución exacta del sistema **A** ▪ **x** = **b** con A simétrica y positiva definida, aunque normalmente lo suele hacer antes. Por otra parte, podemos ver que este método necesita muy poca memoria para llevarse a cabo, lo que le convierte en un gran método cuando la matriz del sistema es simétrica y positiva definida.

Métodos semiempíricos

Existen limitaciones computacionales en el tratamiento de sistemas moleculares con números grandes de electrones. No es suficiente aumentar la memoria ni la velocidad de las computadoras, ni el uso de algoritmos eficientes, los métodos *ab initio* no son rutinariamente aplicados a moléculas con docenas de átomos. Debido a esto se han desarrollado métodos semiempíricos para tratar sistemas con muchos electrones. Los métodos semiempíricos primitivos solo consideran electrones π de moléculas conjugadas. Los métodos semiempíricos son exactos y suficientemente rápidos para ser aplicados rutinariamente a sistemas grandes. Dichos métodos hacen cálculos de estructura electrónica disponibles para un intervalo más amplio de moléculas. No obstante, los métodos *ab initio* representan una aproximación teóricamente más pura. Una de las limitaciones a la exactitud de los métodos semiempíricos, en adición a las aproximaciones inherentes en su formulación, es la exactitud de los datos experimentales usados para obtener los parámetros. Sin embargo, en gran parte debido a que los parámetros ajustables son optimizados para reproducir cierto número de propiedades químicas, los métodos semiempíricos han tenido buena demanda.

Los MO canónicos de una molécula orgánica insaturada plana se pueden dividir en MO σ y π según sean +1 o −1, los eigenvalores para reflexión en el plano molecular, respectivamente. Los métodos semiempíricos primitivos de compuestos orgánicos planares tratan separadamente a los electrones π de los electrones σ. La justificación de la separabilidad $\sigma - \pi$ radica en las diferentes simetrías de los orbitales σ *y* π, y de la mayor polarizabilidad de los electrones π que los hace más susceptibles a perturbaciones como las que ocurren en reacciones químicas. El hamiltoniano π–electrón efectivo H_π está dado por

$$H_\pi = -\frac{\hbar^2}{2m_e}\sum_{i=1}^{n_\pi}\nabla_i^2 + \sum_{i=1}^{n_\pi}\mathrm{V}_i^{\pi,eff} + \frac{1}{2}\sum_{i,j}^{n_\pi}\frac{e^2}{4\pi\varepsilon_0 r_{ij}}$$

donde el primer término es el operador de energía cinética para los π–electrones, $\mathrm{V}_i^{\pi,eff}$ es la energía potencial efectiva para el π–electrón *i* que resulta del campo de potencial del núcleo y todos los $\sigma-$electrones, y el término final representa la energía potencial repulsiva debida a las interacciones $\pi-$interelectrónicas. El hamiltoniano h_i^{π} del núcleo para el $\pi-$electrón *i* está definido por

$$h_i^{\pi} = -\frac{\hbar^2}{2m_e}\nabla_i^2 + \mathrm{V}_i^{\pi,eff}$$

Entonces

$$H_{\pi} = \sum_{i=1}^{n_{\pi}} h_i^{\pi} + \frac{1}{2}\sum_{i,j}^{n_{\pi}} \frac{e^2}{4\pi\varepsilon_0 r_{ij}}$$

El uso de una forma aproximada para el hamiltoniano es característica de métodos semiempíricos. Para encontrar una función de onda $\pi-$electrón ψ_{π} se aplica el principio variacional y, a su vez, se minimiza la integral variacional $\int \psi_{\pi}^{*} H_{\pi} \psi_{\pi}\, d\tau$ para conseguir una energía $\pi-$electrón E_{π}. Las principales teorías $\pi-$electrón de MO son el método MO de electrón-libre, el método MO de Hückel, y el método Pariser-Parr-Pople.

Método MO de electrón-libre (FEMO)

La teoría semiempírica π–electrónica más simple es el método FEMO, en el cual se ignora la repulsión interelectrónica $\frac{e^2}{4\pi\varepsilon_0 r_{ij}}$, y el efecto de los electrones σ se representa por una partícula en una caja cuya función de energía potencial es: $V = 0$ en el interior de la caja, y $V = \infty$ en el exterior de la caja. Entonces, la ecuación de Schrödinger es

$$H_{\pi}\psi_{\pi} = E_{\pi}\psi_{\pi}$$

$$\psi_{\pi} = \prod_{i=j}^{n_{\pi}} \phi_i$$

y aproximando H_{π} como la suma de los $H_{\pi}^{core}(i)$'s, y las energías E_{π} como la suma de las e_i's, se tiene que

$$H_{\pi}^{core}(i)\phi_i = e_i\phi_i$$

Al resolver esta ecuación de eigenvalores para cadenas de moléculas conjugadas, la aproximación FEMO toma la caja en movimiento unidimensional de los electrones, y se encuentra que

$$e_i = \frac{n_i^2 h^2}{8 m_e l^2}, \qquad n_i = 1,2,\ldots,$$

Las n_i's están restringidas por el principio de Pauli: no más que dos electrones pueden estar en un MO espacial dado. En el estado electrónico más bajo, los electrones$-\pi$ libres llenan los $\frac{1}{2}n_\pi$ de los MO más bajos.

Método MO de Huckel (HMO)

El método Hückel es el método de aproximación más simple de la teoría de orbitales moleculares. Su aplicación se restringe al tratamiento de sistemas de hidrocarburos planos con enlaces π conjugados como el etano, benceno, butadieno, etc. La aproximación de Hückel asume que los electrones en los enlaces π *sienten* un potencial electrostático debido al enlace σ en la molécula (es decir, se enfoca solo en la formación de enlaces π, dado que el enlace σ ya se ha formado). La teoría de los orbitales moleculares (MO) utiliza una combinación lineal de orbitales atómicos (LCAO) para representar los orbitales moleculares que resultan de los enlaces entre átomos. Estos se dividen en tres tipos, enlazantes, antienlazantes y no-enlazantes.

Pasos para derivar los diagramas de MO

- Encuentre la configuración de electrones de valencia de cada átomo del sistema.
- Decida si el sistema es homonuclear o heteronuclear.
- Llene los MO usando la energía y las propiedades de enlace de los orbitales atómicos superpuestos.
- Use el diagrama para predecir las propiedades del sistema

Método Pariser-Parr-Pople (PPP)

El método PPP considera solo los orbitales π de las moléculas, y es una mejora con respecto a la teoría HMO. Tanto el método PPP como CNDO se basan en la aproximación ZDO. La

integral de Coulomb de dos electrones se trata como un parámetro empírico y se ignoran varias integrales de dos electrones (Gupta, 2016).

Mecánica molecular

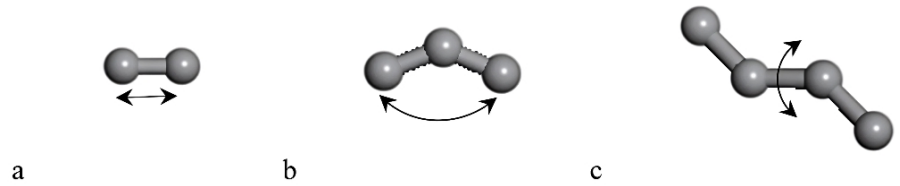
La mecánica molecular calcula la energía de una molécula ajustándola con cambios en longitudes y ángulos de enlace hasta tener una estructura de mínima energía. La energía estérica de una molécula se debe a las energías de enlace y térmica. La mecánica molecular calcula la energía estérica de una molécula debida a su geometría o conformación. Entonces, lo que tiene significado físico en mecánica molecular es la diferencia de energía-estérica (calculada usando el mismo campo de fuerza) entre dos especies que tienen el mismo número y tipo de átomos y el mismo número y tipo de enlaces, por ejemplo: a) diferentes conformadores de la misma molécula, b) diferentes estereoisómeros de la molécula, c) diferentes especies por rotación de un enlace, d) diferentes geometrías de la misma molécula, e) dos moléculas alejadas entre sí y el mismo par formando un enlace de hidrógeno. La diferencia entre las energías estéricas de equilibrio de dos especies con los mismos números y tipos de enlaces es una estimación de la diferencia entre las energías electrónicas de equilibrio de dos especies. Correcciones a las diferencias de energía vibracionales y diferencias de energía térmicas deberían agregarse a la diferencia de energía estérica, pero son usualmente pequeñas y frecuentemente se omiten. La energía estérica de mecánica-molecular se puede combinar con parámetros de energía-de-enlace empíricos para calcular calores de formación $\Delta H^o_{f,298}$. Tomando $\sum_i Q_i \boldsymbol{r}_i$ donde la sumatoria va sobre todas las cargas parciales atómicas, uno puede calcular el momento dipolar molecular para una conformación dada. Evaluando las segundas derivadas parciales de la energía en un mínimo local se pueden calcular las frecuencias vibracionales moleculares. Una optimización de geometría en mecánica molecular empieza con una geometría inicialmente supuesta, y encuentra la energía estérica (energía debida a la geometría o conformación de la molécula) mínima local más cercana, por minimización de la energía usando algún algoritmo de optimización.

Introducción a campos de fuerza (*force fields*)

Para cualquier cálculo molecular se debe construir un modelo que se aproxime a la realidad. El modelo para un cálculo de mecánica molecular es que la energía de una molécula se

describa en términos de una función conocida como campo de fuerza (*force field*), que depende de las posiciones atómicas. Esta función debe proporcionar una descripción de las fuerzas que actúan dentro de la molécula. Para moléculas orgánicas, una de esas funciones es el campo de fuerza de valencia, mediante el cual se determina la energía de la molécula en términos de coordenadas internas, longitudes de enlace, ángulos de enlace, ángulos diedrales y coordenadas cartesianas de los átomos. Esto se ilustra esquemáticamente en la figura 37. El término campo de fuerza se toma para referirse a un campo de fuerza de valencia.

Figura 37. *Diagrama esquemático de los términos de campo de fuerza*



Nota. a) Estiramiento de enlace, b) flexión del ángulo de enlace y c) rotación diedral.
Fuente: elaboración propia.

Hay muchos campos de fuerza diferentes que usan formas para distintas interacciones en y entre las moléculas. La forma particular de un campo de fuerza depende de la precisión requerida para su propósito. Los parámetros de campo de fuerza se encuentran a partir de datos experimentales o teóricos. Una propiedad clave tanto para la forma del campo de fuerza como para sus parámetros es que aquellos para un átomo o grupo de átomos deben ser los mismos para diferentes moléculas, es decir, deben ser *transferibles*. Sin esta propiedad, habría que construir un campo de fuerza para cada molécula. Esta noción de transferibilidad aparece en gran parte de la química y se observa que se mantiene en muchos casos. Por ejemplo, las frecuencias de estiramiento de los enlaces carbono-hidrógeno son en gran medida independientes del entorno molecular, al igual que las longitudes de los enlaces C-H. Subdividir estos en enlaces entre el hidrógeno y los átomos de carbono con diferentes hibridaciones conduce a un acuerdo más fuerte.

Si una molécula fuera una colección de bolas (átomos) conectadas por resortes (enlaces), su energía potencial podría describirse conociendo la masa de todas las bolas y las constantes de resorte de todos los resortes. Computar el arreglo correspondiente de las

energías más bajas (minimización de la energía) indicaría la geometría probable del sistema. También se podría calcular la conducta dinámica en respuesta a cualquier perturbación, aplicando las leyes de la mecánica clásica.

En realidad, los sistemas moleculares son mucho más complejos que este simple esquema, pero el principio de mecánica molecular es similar. El sistema se describe por una expresión que suma todos los componentes energéticos. En adición a la energía de enlace, que puede describirse como un *resorte* en la manera antes bosquejada, esta incluye términos como la energía de Van der Waals, la energía de Coulomb (para sistemas cargados), la energía requerida para retorcer torsiones y distorsionar ángulos de enlace, y así sucesivamente.

$$E_{total} = E_{enlace} + E_{Coulomb} + E_{VdW} + E_{ángulo} + E_{torsión} + \cdots$$

Esta expresión de energía es la base de un *force field* debido a que habilita la energía, y en consecuencia las fuerzas del sistema, a ser calculado. Nótese que al hacer cálculos de punto simple (*single point*) para obtener curvas de energía potencial, la velocidad de los átomos o moléculas en consideración es cero; por lo tanto, la energía cinética es cero. En consecuencia, la energía total es igual a la energía potencial.

i) Energía de estiramiento del enlace

La energía E_{enlace} de estiramiento del enlace es la suma de las energías por pares del estiramiento del enlace en cada par de átomos enlazados, eligiendo además la aproximación más simple que es la del oscilador armónico. Entonces

$$E_{enlace}(r_{ij}) = \frac{k_r}{2}(r_{ij} - r_0)^2$$

donde r_{ij} es la distancia entre los átomos i y j, r_0 es su distancia de equilibrio y k_r es la constante de enlace. Los parámetros r_0 y k_r se pueden obtener por técnicas experimentales, o por métodos teóricos.

ii) Energía de Coulomb

La energía de Coulomb $E_{Coulomb}$ entre dos cargas, o energía potencial electrostática en sistemas moleculares, también se puede expresar como

$$E_{Coulomb} = \frac{1}{4\pi\epsilon_0}\frac{q_{ia}q_{jb}}{r_{iajb}}$$

Siendo ϵ_0 la permitividad en el vacío y r_{iajb} la distancia relativa entre la carga q_{ia} del átomo a en la molécula i, y q_{jb} la carga del átomo b en la molécula j.

iii) *Energía de Van der Waals*

La energía de Van der Waals E_{VdW} usualmente se toma como la suma de las interacciones que involucran todas las interacciones posibles de pares de átomos:

$$E_{VdW,ij} = \varepsilon_{ij}\left[\left(\frac{r_{ij}^0}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^0}{r_{ij}}\right)^{6}\right] = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

donde r_{ij} es la distancia entre los átomos i y j, ε_{ij} es la profundidad del pozo en el mínimo de la curva de energía potencial, el parámetro r_{ij}^0 da el valor de r_{ij} en el mínimo de $E_{VdW,ij}$ y el parámetro σ_{ij} es la distancia en la cual $E_{VdW,ij} = 0$.

iv) Energía de ángulo

La energía de ángulo $E_{ángulo}$ es la suma de las energías de todos los ángulos entre ternas de átomos en la molécula. Estas se representan por una función de tipo armónico

$$E_{ángulo}(\theta_{ijk}) = \frac{k_\theta}{2}(\theta_{ijk} - \theta_0)^2$$

donde θ_{ijk} es el ángulo de enlace entre los átomos i, j, k; θ_0 es el ángulo de equilibrio y k_θ es la constante de enlace que se puede obtener experimental o teóricamente.

v) Energía de torsión

La energía de torsión pertenece a las energías de ángulo de torsión, las cuales corresponden a los ángulos de torsión propios e impropios. La forma más simple asociada a ángulos de torsión propios es una función periódica que involucre a los cuatro átomos (i, j, k, l)

$$E_{torsión\ propia}(\phi_{ijkl}) = \sum_n C_n cos^n(\phi_{ijkl})$$

donde ϕ_{ijkl} es el ángulo de torsión entre los cuatro átomos (i, j, k, l), las constantes C_n de las torsiones se ajustan con cálculos *ab initio* de superficies de potencial. Los ángulos de torsión impropios se representan por una función de tipo armónico de la siguiente manera

$$E_{torsión\ impropia}(\phi_{ijkl}) = \frac{k_\phi}{2}(\phi_{ijkl} - \phi_0)^2$$

Donde k_ϕ es la constante de enlace y ϕ_0 es la constante de equilibrio.

COMPASS como modelo de campo de fuerza (*force field*)

Para describir diferentes sistemas se usan modelos diferentes en el force field COMPASS. Para sistemas moleculares covalentes se usa el modelo de valencia CFF, que es el modelo más extensivamente parametrizado en COMPASS. Para sistemas iónicos (haluros alcalinos) se usa el modelo iónico. Se usa una función VdW (LJ-9-6) simple para describir las propiedades de bulto de metales. Para metales y transiciones de óxido de metal se introduce un modelo semiiónico. La energía total se escribe como una combinación de términos de valencia, que incluyen los términos de acoplamiento cruzado en la diagonal y fuera de diagonal, y términos de interacción no enlazantes. Los términos de valencia contienen $E_b, E_\theta, E_\phi y\ E_\chi$ para enlace, ángulo, torsión y coordenadas de ángulo fuera-del-plano, respectivamente, y $E_{bb'}, E_{b\theta}, E_{b\phi}, E_{\theta\theta'}, y\ E_{\theta\theta'\phi}$ para términos de acoplamiento cruzado entre las coordenadas internas. Su forma funcional es la siguiente (Sun, 1998):

$$E_{total} = E_{enlazante} + E_{noenlazante}$$

$$E_{enlazante} = E_b + E_\theta + E_\phi + E_\chi E_{bb'} + E_{b\theta} + E_{b\phi} + E_{\theta\theta'} + E_{\theta\theta'\phi}$$

$$E_{noenlazante} = E_{elec} + E_{LJ}$$

$$E_{total} = E_b + E_\theta + E_\phi + E_\chi + E_{bb'} + E_{b\theta} + E_{b\phi} + E_{\theta\theta'} + E_{\theta\theta'\phi}$$

$$+E_{elec} + E_{LJ}$$

$$= + \sum_b [k_2(b - b_0)^2 + k_3(b - b_0)^3 + k_4(b - b_0)^4]$$

$$+\sum_{\theta}[H_2(\theta-\theta_0)^2+H_3(\theta-\theta_0)^3+H_4(\theta-\theta_0)^4]$$

$$+\sum_{\phi}[V_1(1-cos\phi)+V_2(1-cos2\phi)+V_3(1-cos3\phi)]$$

$$+\sum_{\chi}K_{\chi}(\chi-\chi_0)^2+\sum_{b,b'}F_{b,b'}(b-b_0)(b'-b_0')$$

$$+\sum_{b,\theta}F_{b\theta}(b-b_0)(\theta-\theta_0)+\sum_{b,\phi}(b-b_0)$$

$$\left[F_{b,\phi}^{(1)}(1-cos\phi)+F_{b,\phi}^{(2)}(1-2cos\phi)+F_{b,\phi}^{(3)}(1-cos3\phi)\right]$$

$$+\sum_{\theta,\theta'}F_{\theta,\theta'}(\theta-\theta_0)(\theta'-\theta_0')$$

$$+\sum_{\theta,\theta'\phi}F_{\theta,\theta'\phi}(\theta-\theta_0)(\theta'-\theta_0')cos\phi$$

$$+\sum_{i,j}\frac{q_iq_j}{\varepsilon_0 r_{ij}}+\sum_{i,j}\varepsilon_{ij}\left[2\left(\frac{r_{ij}^0}{r_{ij}}\right)^9-3\left(\frac{r_{ij}^0}{r_{ij}}\right)^6\right]$$

Modelo covalente

Para todas las moléculas orgánicas e inorgánicas de enlace covalente, incluyendo polímeros, se usa la forma funcional CFF. La energía total se escribe como una combinación de términos de valencia que incluyen los términos de acoplamiento cruzado en la diagonal y fuera de diagonal:

$$E_{valencia}=\sum_{b}[k_2(b-b_0)^2+k_3(b-b_0)^3+k_4(b-b_0)^4]$$

$$+\sum_{\theta}[k_2(\theta-\theta_0)^2+k_3(\theta-\theta_0)^3+k_4(\theta-\theta_0)^4]$$

$$+\sum_{\phi}[k_1(1-cos\phi)+k_2(1-2cos\phi)+k_3(1-cos3\phi)]$$

$$+\sum_{\chi}k_2(\chi-\chi_0)^2+\sum_{b,b'}k(b-b_0)(b'-b_0')$$

$$+\sum_{b,\theta}k(b-b_0)(\theta-\theta_0)$$

$$+\sum_{b,\phi} k(b-b_0)\,(1-cos\phi)$$

$$+\sum_{\theta,\theta'\phi} F_{\theta,\theta'\phi}(\theta-\theta_0)(\theta'-\theta_0')cos\phi$$

más términos no enlazantes que incluyen la energía coulómbica:

$$E_{elec}=\sum_{i,j}\frac{q_i q_j}{\varepsilon_0 r_{ij}}$$

La interacción electrostática se representa por el modelo de carga atómica usando incrementos de enlace, lo cual representa la separación de carga entre dos átomos enlazados. La carga parcial neta para el átomo *i* (denotada q_i) es la suma de todos los incrementos de enlace sobre todos los átomos enlazados a este átomo:

$$q_i=\sum_j \delta_{ij}$$

La interacción VdW usa la función de Lenard-Jones-9-6

$$E_{VdW}=\sum_{i,j}\varepsilon_{ij}\left[2\left(\frac{r_{ij}^0}{r_{ij}}\right)^9-3\left(\frac{r_{ij}^0}{r_{ij}}\right)^6\right]$$

Los términos no-enlazantes se usan para interacciones entre pares de átomos que están separados por dos o más átomos interviniendo, que pertenecen a diferentes moléculas.

Los parámetros de Lennard-Jones-9-6 son para átomos pares. Para átomos diferentes de pares se usa una combinación de reglas de sexto orden para los parámetros fuera-de-la-diagonal:

$$r_{ij}^0=\left(\frac{(r_i^0)^6+\left(r_j^0\right)^6}{2}\right)^{1/6}$$

$$\varepsilon_{i,j}=2\sqrt{\varepsilon_i\cdot\varepsilon_j}\left(\frac{(r_i^0)^3\cdot\left(r_j^0\right)^3}{\left(r_i^0\right)^6\cdot\left(r_j^0\right)^6}\right)$$

Modelo iónico

Un modelo iónico simple únicamente con los términos coulómbico y vdW se utiliza en COMPASS. El uso de las mismas formas funcionales asegura compatibilidad de los

parámetros cuando se mezclan diferentes sistemas. En este modelo, cada átomo es tratado como una partícula no enlazada, no hay un *enlace* de valencia entre ningún par de átomos. El enlace iónico está representado por una fuerte fuerza de atracción entre las especies con carga opuesta y los términos vdW cuya parte de repulsión mantiene a las especies con carga opuesta a cierta distancia. Las cargas –q_s– son cargas formales que en la mayoría de los casos son equivalentes a los estados de oxidación. Los incrementos de enlace no se usan en este modelo.

Modelo semiiónico para óxidos metálicos

Se utiliza un modelo semiiónico para especies que están entre lo puramente iónico y lo covalente. Este modelo, similar al modelo iónico, trata a cada átomo como *no enlazado*. Los mismos términos no enlazados se utilizan como predeterminados para cualquier par de átomos que no estén inmediatamente en contacto. Para los átomos que están en contacto directo (enlazados), la energía de interacción se representa mediante un término electrostático y una función de dispersión de Morse:

$$E_{M-D} = \varepsilon\left\{e^{[-2\alpha(r-r_m)]} - 2e^{[-\alpha(r-r_m)]}\right\} \cdot f_s + \frac{C_6}{r^6} \cdot (1 - f_s)$$

$$con \qquad f_s = \frac{1}{1 + e^{(r-r_c)}}$$

f_s es una función continua conmutante que transfiere la función de Morse de corto alcance a una función de dispersión de largo alcance.

Los parámetros (q) de carga real, que pueden ser determinados de datos *ab initio*, dependen del elemento y su estado de oxidación. Generalmente hablando son cargas parciales. Las contribuciones del término electrostático y la función de dispersión-de-Morse representan las contribuciones iónica y covalente al enlace, respectivamente. Los términos vdW restantes representan las interacciones no-enlazantes de largo-alcance. Por lo tanto, los parámetros vdW son consistentes con aquellos que se usan en los modelos covalente y iónico. Es razonable suponer que las combinaciones de los parámetros vdW entre diferentes átomos en sistemas mezclados es una buena aproximación para las interacciones no-enlazantes entre diferentes sistemas.

Modelo de Lennard-Jones para metales

Usar un método de campo de fuerzas para estudiar metales y aleaciones es una tarea difícil. Como una primera aproximación, en COMPASS se usa un modelo muy crudo —la función LJ-9-6— para describir metales puros en el bulto. Las propiedades usadas para derivar los parámetros son energías de red y estructuras, basados en cálculos de minimización de la energía. Ya que los parámetros representan las contribuciones totales efectivas tanto de enlaces metálicos como interacciones de vdW de largo-alcance, los parámetros podrían no ser adecuados al mezclarlos con otros parámetros de vdW para representar interacciones no-enlazantes.

Capítulo 3. Dinámica molecular

Se conoce como dinámica molecular a la solución computacional de las ecuaciones de movimiento de sistemas de partículas, con aproximaciones como la de diferencias finitas. Dicha solución depende de una descripción del movimiento de partículas, que exprese la distancia, la velocidad y la aceleración en serie de Taylor, tal como lo hace el método de Verlet. Un problema sustancial es elegir un potencial intermolecular del sistema en estudio. Uno de los más usados es el potencial de Lennard-Jones.

Ecuaciones de movimiento para sistemas atómicos

Las ecuaciones de movimiento se pueden establecer al menos de tres formas: newtoniana, lagrangiana y hamiltoniana. Estos son tres sistemas consistentes que se aplican para encontrar y resolver analíticamente las ecuaciones de movimiento de problemas de dinámica de partículas. Cuando las ecuaciones de movimiento resultan de mayor complejidad se aprovechan los métodos numéricos para encontrar su solución.

Tres modelos concordantes

a) Modelo newtoniano

$$F_i = m_i \ddot{r}_i$$

F_i, r_i , y m_i son la fuerza, la distancia y la masa del átomo i. La fuerza y la distancia son cantidades vectoriales, mientras que la masa es escalar.

$$F_i = -\nabla_{r_i} V$$

Las coordenadas del sistema se obtienen con el diagrama de equilibrio.

b) Modelo lagrangiano

$$L(q, \dot{q}) = T - V$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0$$

La función lagrangiana L está en función de las coordenadas generalizadas q_k y sus derivadas con respecto al tiempo $\dot{q}_k$. Las ecuaciones de movimiento se consideran aplicadas al centro de masa del movimiento de una molécula.

c) Modelo hamiltoniano

Las posiciones y los momentos son cantidades vectoriales. El momento generalizado p_k conjugado a q_k es

$$p_k = \frac{\partial L}{\partial \dot{q}_k}$$

La forma hamiltoniana de las ecuaciones de movimiento está representada por las ecuaciones diferenciales de segundo orden

$$\dot{q}_k = \frac{\partial H}{\partial p_k}$$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}$$

Las cuales se expresan como ecuaciones de primer orden como sigue

$$\dot{r}_i = \frac{p_i}{m}$$

$$\dot{p}_i = -\nabla_{r_i} V$$

$$= F_i$$

El hamiltoniano está definido por la ecuación

$$H(p,q) = \sum_k \dot{q}_k p_k - L(q,\dot{q})$$

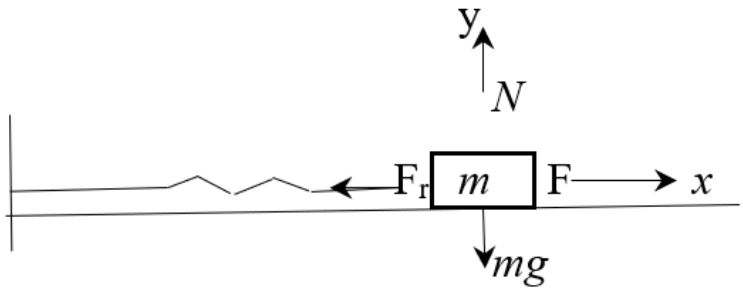
$\dot{q}_k$ se puede escribir como una función del momento.

Se computan las trayectorias del centro de masa de las ecuaciones de Hamilton, lo que hace que se tenga que resolver un sistema de 3N ecuaciones diferenciales de segundo orden, o un conjunto equivalente de 6N ecuaciones diferenciales de primer orden.

Ejemplo:

Para ejemplificar, a continuación se muestran las ecuaciones de movimiento de un oscilador armónico amortiguado, en cada uno de estos tres sistemas: una masa m está sometida a una fuerza restauradora lineal –kx, y a una fuerza amortiguadora μmg causada por la fricción al deslizarse (ver figura 38). Se le da un desplazamiento inicial A y luego se suelta. Halle el movimiento.

Figura 38. *Oscilador armónico amortiguado*



Fuente: elaboración propia.

Formulación newtoniana

$$\sum F_x = -F_r - F_a = 0, \quad \sum F_y = N - mg = 0$$

Ya que la fuerza del resorte y la fuerza de fricción se han puesto en la misma dirección,

$$\sum_i F_i = 0 \Rightarrow F - F_r - F_a = 0 \quad \text{donde} \quad F_a = f = \mu mg, \quad F_r = -kx$$

Por lo tanto $F = F_r + F_a \Rightarrow ma = -kx + \mu mg$ y la ecuación de movimiento es

$$m\ddot{x} = -kx + \mu mg$$

Formulación lagrangiana

$$L = T - V$$

$$T = \frac{1}{2}m\dot{x}^2, \quad V = -\int F \cdot dr \quad = -\int(-kx + \mu mg)dx$$

$$= \tfrac{1}{2}kx^2 - \mu mg$$

$$L = \frac{1}{2}m\dot{x}^2 - \left(\frac{1}{2}kx^2 - \mu mg\right)$$

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 + \mu mg$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0 \Rightarrow \frac{d}{dt}(m\dot{x}) + kx - \mu mg = 0$$

$$\therefore m\ddot{x} = -kx + \mu mg$$

Formulación hamiltoniana

$$p_k = \frac{\partial L}{\partial \dot{q}_k} \qquad \Rightarrow \qquad p = m\dot{x}$$

$$H(p,q) = \sum_k \dot{q}_k p_k - L(q,\dot{q}) \qquad \Rightarrow \qquad H = \dot{x}p - \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 - \mu mg$$

$$H = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}kx^2 - \mu mg$$

$$\dot{q}_k = \frac{\partial H}{\partial p_k}$$

$$\dot{p}_k = \frac{\partial H}{\partial q_k} \quad \Rightarrow \quad \begin{cases} \dot{p} = -kx + \mu mg \\ \therefore \ m\ddot{x} = -kx + \mu mg \end{cases}$$

La solución a la ecuación diferencial $m\ddot{x} = -kx + \mu mg$ con movimiento hacia la izquierda, obtenida de tres maneras distintas correspondientes a los métodos Newtoniano, Lagrangiano y Hamiltoniano, es la siguiente:

$$\ddot{x} + \omega_0^2 x = \mu g \qquad con \qquad \omega_0^2 = k/m$$

$$r^2 + \omega_0^2 = 0 \qquad \rightarrow \qquad r = \pm\omega_0 i$$

La solución homogénea es: $x_H = C_1 cos\omega_0 t + C_2 sen\omega_0 t$

Y la solución particular es: $x_P = C_3$

al sustituir en la ecuación diferencial se tiene que $C_3 = \mu g/\omega_0^2$. Y como $x = x_H + x_P \Rightarrow$

$$x = C_1 cos\omega_0 t + C_2 sen\omega_0 t + \mu g/\omega_0^2$$

$$\dot{x} = -\omega_0 C_1 sen\omega_0 t + \omega_0 C_2 cos\omega_0 t$$

Como el desplazamiento inicial es A, entonces

$$C_1 = -\mu g/\omega_0^2 + A$$

$$C_2 = 0$$

$$\therefore \quad x(t) = (A - \mu g/\omega_0^2) cos\omega_0 t + \mu g/\omega_0^2$$

El máximo desplazamiento hacia la izquierda se obtiene cuando $\omega_0 t = \pi$. Por lo tanto

$$x_1 = -(A - \mu g/\omega_0^2) + \mu g/\omega_0^2 \ \ o\ bien\ x_1 = -A + \frac{2\mu g}{\omega_0^2}$$

x_1 es el punto de retorno, y ahí mismo se inicia el movimiento hacia la derecha. La ecuación de movimiento es $\ddot{x} + \omega_0^2 x = -\mu mg$ con condiciones iniciales $x(0) = x_0,\ \dot{x}(0) = 0$.

$$x = C_1 cos\omega_0 t + C_2 sen\omega_0 t - \mu g/\omega_0^2$$

$$\dot{x} = -\omega_0 C_1 sen\omega_0 t + \omega_0 C_2 cos\omega_0 t$$

$$C_1 = x_1 + \mu g/\omega_0^2$$

$$C_2 = 0$$

$$\therefore \quad x(t) = (x_1 + \mu g/\omega_0^2) cos\omega_0 t - \mu g/\omega_0^2$$

En este caso el máximo desplazamiento se obtiene cuando $\omega_0 t = -\pi$ y como $cos(-\pi) = -1$ entonces $x_2 = -(-A + 3\mu g/\omega_0^2) - \mu g/\omega_0^2 \ \ o\ bien\ x_2 = A - \frac{4\mu g}{\omega_0^2}$

Conservación de la energía

Si la energía es constante para cualquier posición de la partícula, entonces la energía de la partícula se conserva.

Si la invariancia traslacional se preserva, entonces el momento lineal total se conserva. Esto ocurre en el caso de condiciones de frontera periódicas.

Si un sistema (o sea, el hamiltoniano *H*) es invariante a traslación en una dirección particular, entonces el momento lineal correspondiente se conserva.

Si el sistema es invariante a rotación alrededor de un eje, entonces la componente del momento angular correspondiente se conserva

Energía potencial

$$V = \sum_i v_1(r_i) + \sum_i \sum_{j>i} v_2(r_i, r_j) + \sum_i \sum_{j>i} \sum_{k>j>i} v_3(r_i, r_j, r_k) + \cdots$$

Ejemplo:

Si un sistema queda encerrado en una caja esférica con el potencial v_1 esféricamente simétrico, entonces los tres componentes del momento angular total alrededor del centro de simetría se conservan, y el momento traslacional total no se conserva.

Métodos de diferencias finitas

Teorema de Taylor

Suponga que $f \epsilon C^n[a,b]$, que $f^{(n+1)}$ existe en $[a,b]$ y $x_0 \epsilon [a,b]$. Para cada $x \epsilon [a,b]$ existe un número $\xi(x)$ entre x_0 y x tal que

$$f(x) = P_n(x) + R_n(x)$$

donde

$$P_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

y

$$R_n(x) = \frac{f^{(n+1)}\big(\xi(x)\big)}{(n+1)!}(x - x_0)^{n+1}$$

$P_n(x)$ es el n-ésimo polinomio de Taylor para f respecto a x_0, y $R_n(x)$ es el término del residuo (o error de truncamiento) asociado a $P_n(x)$. La serie infinita obtenida al tomar el límite de $P_n(x)$ cuando $n \to \infty$ es la serie de Taylor para f en torno a x_0.

Para resolver las ecuaciones de movimiento:

$$m_i \ddot{r}_i = f_i$$

$$\dot{r}_i = p_i/m$$

$$\dot{p}_i = -\nabla_{r_i} V = F_i$$

se utilizan diferencias finitas. Esto es, dadas las posiciones moleculares, velocidades y otra información dinámica en el tiempo, se trata de obtener posiciones, velocidades, etcétera en un tiempo posterior $t + \delta t$ a un grado de exactitud suficiente. Si la trayectoria clásica es continua, entonces una estimación de posiciones, velocidades, etc. en el tiempo $t + \delta t$ se puede obtener por una expansión en serie de Taylor alrededor del tiempo t:

$$r(t + \delta t) = r(t) + \dot{r}\delta t + \frac{1}{2!}\ddot{r}(t)\delta t^2 + \frac{1}{3!}\dddot{r}(t)\delta t^3 + \cdots$$

$$r^p(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) + \frac{1}{6}\delta t^3 b(t) + \cdots$$

$$v^p(t + \delta t) = v(t) + \delta t a(t) + \frac{1}{2}\delta t^2 b(t) + \cdots$$

$$a^p(t + \delta t) = a(t) + \delta t b(t) + \cdots$$

$$b^p(t + \delta t) = b(t) + \cdots$$

Con estas ecuaciones no se generan trayectorias correctas conforme avanza el tiempo, debido a que no se han introducido las ecuaciones de movimiento. Se necesita un paso de corrección. De las nuevas posiciones r^p se pueden calcular las fuerzas en el tiempo $t + \delta t$, y en consecuencia las aceleraciones correctas $a^c(t + \delta t)$. Estas se pueden comparar con la aceleración predicha, para estimar el tamaño del error en el paso de predicción en la forma

$$\Delta a(t + \delta t) = a^c(t + \delta t) - a^p(t + \delta t)$$

Este error, y los resultados del paso predictor, se alimentan en el paso corrector, que se lee de la siguiente manera

$$r^c(t+\delta t) = r^p(t+\delta t) + C_0 \Delta a(t+\delta t)$$

$$v^c(t+\delta t) = v^p(t+\delta t) + C_1 \Delta a(t+\delta t)$$

$$a^c(t+\delta t) = a^p(t+\delta t) + C_2 \Delta a(t+\delta t)$$

$$b^c(t+\delta t) = b^p(t+\delta t) + C_3 \Delta a(t+\delta t)$$

El esquema general de una simulación DM (dinámica molecular) paso a paso basada en el algoritmo predictor-corrector es el siguiente:

a) Predecir las posiciones, velocidades, aceleraciones, etcétera a un tiempo $t + \delta t$, usando los valores comunes de estas cantidades.
b) Evaluar las fuerzas y, en consecuencia, las aceleraciones $a_i = f_i/m$ de las nuevas posiciones.
c) Corregir las posiciones predichas, velocidades, aceleraciones, etcétera, usando las nuevas aceleraciones.
d) Calcular cualquier variable de interés, tal como la energía, el coeficiente del virial, parámetros de orden, listas para la acumulación de promedios temporales, antes de regresar a (a) para el siguiente paso.

La exactitud y estabilidad de un algoritmo de simulación se miden por sus errores de truncamiento local y global, y el algoritmo se puede probar en un modelo simple, tal como el oscilador armónico.

Algoritmo de Verlet

Es una solución numérica de ecuaciones de movimiento de segundo orden. Dicha solución se lleva a cabo con métodos como el de Euler o Runge-Kutta. Dado que se debe plantear un sistema de ecuaciones, el número de operaciones aumenta al realizar su implementación en un programa computacional, de modo que el algoritmo de Verlet (por trabajar con la segunda derivada) tiene ventajas en aplicaciones en que interviene la mecánica clásica. El algoritmo de Verlet se deduce a partir de la expansión en serie de Taylor de la función de posición $r(t)$ en $r(t+\delta t)$.

$$r(t+\delta t) = r(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) + \cdots$$

y de la expansión en serie de Taylor 'hacia atrás':

$$r(t-\delta t) = r(t) - \delta t v(t) + \frac{1}{2}\delta t^2 a(t) + \cdots$$

Sumando estas dos ecuaciones se puede observar que el método está basado en posiciones $r(t)$, y aceleraciones $a(t)$, y las posiciones del paso previo $r(t-\delta t)$. Por lo que la ecuación para avanzar posiciones es:

$$r(t+\delta t) = 2r(t) - r(t-\delta t) + \delta t^2 a(t) + \cdots$$

Las velocidades se eliminaron por adición de las ecuaciones obtenidas por la expansión en serie de Taylor alrededor de $r(t)$, tal que restándolas ahora se obtienen velocidades útiles para calcular la energía cinética, se obtienen diferencias centradas:

$$v(t) = \frac{r(t+\delta t) - r(t-\delta t)}{2\delta t}$$

Regresando a la ecuación previa a esta última, la segunda ley de Newton se puede expresar iterativamente como

$$r_{n+1} = 2r_n - r_{n-1} + \frac{h^2}{m} F(r_n, t_n)$$

En la cual se ha representado el incremento en el tiempo δt por el tamaño de paso h. Para su inicialización, además se debe calcular el valor correspondiente a $r(-h)$, que se obtiene del desarrollo previo en serie de Taylor 'hacia atrás':

$$r_{n-1} = r_n - hv_n + \frac{h^2}{2m} F(r_n, t_n)$$

La ecuación para avanzar posiciones es correcta a errores del orden de δt^4, mientras que la ecuación para las velocidades está sujeta a errores de orden δt^2.

El cálculo de la energía cinética y el momento lineal se deben incluir en un *loop* dentro de un algoritmo adecuado, puesto que es la única vez que $r(t+\delta t)$ y $r(t-\delta t)$ están disponibles para calcular velocidades. Se pueden evaluar las fuerzas del siguiente paso siguiendo el movimiento de la partícula.

Esquema salto de rana (*leap-frog*)

Al expresar las velocidades en ecuaciones de diferencias se tiene que

$$v\left(t+\frac{\delta t}{2}\right) \equiv \frac{r(t+\delta t)-r(t)}{\delta t}$$

$$v\left(t-\frac{\delta t}{2}\right) \equiv \frac{r(t)-r(t-\delta t)}{\delta t}$$

Se resta

$$v\left(t+\frac{\delta t}{2}\right)-v\left(t-\frac{\delta t}{2}\right)=\frac{-2r(t)+r(t-\delta t)+r(t+\delta t)}{\delta t}$$

Del lado derecho de esta igualdad resulta la ecuación en diferencias δt veces la segunda derivada con respecto al tiempo, que en términos de la aceleración se expresa como

$$v\left(t+\frac{\delta t}{2}\right)-v\left(t-\frac{\delta t}{2}\right)=a(t)\delta t$$

De tal modo que las ecuaciones *leap and frog* son:

$$r(t+\delta t)=r(t)+\delta t v\left(t+\frac{1}{2}\delta t\right)$$

$$v\left(t+\frac{1}{2}\delta t\right)=v\left(t-\frac{1}{2}\delta t\right)+\delta t a(t).$$

Las cantidades que se almacenan son posiciones comunes $r(t)$ y aceleraciones $a(t)$ junto con velocidades a mitad de paso $v\left(t-\frac{1}{2}\delta t\right)$. Esta ecuación de la velocidad se implementa primero, y las velocidades saltan (leap) sobre las coordenadas para dar el valor de medio paso $v\left(t+\frac{1}{2}\delta t\right)$. Durante este paso las velocidades se calculan como

$$v(t)=\frac{1}{2}\left(v\left(t+\frac{1}{2}\delta t\right)-v\left(t-\frac{1}{2}\delta t\right)\right)$$

El método de salto de rana (*leap-frog*) aun no maneja las velocidades de una manera satisfactoria. Un algoritmo Verlet-equivalente que almacene posiciones, velocidades y aceleraciones al mismo tiempo t, y que minimice el error de redondeo se consigue en el de *velocidad de Verlet*:

$$r(t+\delta t) = r(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t)$$

$$v(t+\delta t) = v(t) + \frac{1}{2}\delta t[a(t) + a(t+\delta t)]$$

La primera de estas últimas dos ecuaciones se deduce de las anteriores, recordando la derivada como ecuación de diferencias, es decir:

$$v(t) = \frac{r(t) - r(t-\delta t)}{\delta t}$$

La segunda ecuación se deduce al aplicar la regla del trapecio, esto es

$$v(t+\delta t) = v(t) + \int_t^{t+\delta t} a(t)dt$$

El algoritmo de Verlet se recupera eliminando las velocidades. En esta forma el método asemeja un algoritmo predictor-corrector de tres valores, donde el coeficiente corrector de posición es cero. El algoritmo solo requiere almacenar $r, v,$ y a. Las nuevas posiciones al tiempo $t + \delta t$ se calculan con la ecuación anterior de la velocidad, y las velocidades en el paso medio se calculan como sigue

$$v\left(t + \frac{1}{2}\delta t\right) = v(t) + \frac{1}{2}\delta t a(t)$$

Con esto ya se pueden calcular las fuerzas y aceleraciones al tiempo $t + \delta t$, y la velocidad se mueve a:

$$v(t+\delta t) = v\left(t + \frac{1}{2}\delta t\right) + \frac{1}{2}\delta t a(t+\delta t)$$

En este punto está disponible la energía cinética al tiempo $t + \delta t$. La energía potencial a este tiempo habrá sido evaluada en el *loop* de fuerza (Allen y Tildesley, 1987).

Dinámica molecular de cuerpos rígidos no-esféricos

Los movimientos de interés son el traslacional y el rotacional del centro de masa que ocurren en la mecánica clásica. El primero se maneja por los métodos previamente mencionados: se

interpreta la fuerza f_i en la ecuación $m_i\ddot{r}_i = f_i$ como la suma vectorial de todas las fuerzas que actúan en la molécula i en el centro de masa r_i. El movimiento rotacional se gobierna por la torca τ_i alrededor del centro de masa. Cuando las interacciones tienen la forma de fuerzas f_{ia} que actúan en sitios r_{ia} en la molécula, la torca se define por

$$\tau_i = \sum_a (r_{ia} - r_i) \times f_{ia} = \sum_a d_{ia} \times f_{ia}$$

las posiciones de átomos relativas al centro de masa molecular se definen como d_{ia}. Cuando aparecen términos multipolares en el potencial las fuerzas y las torcas provienen de tres modelos de potencial posibles:

i) Una cadena polimérica con longitudes de enlace restringidas, pero inclinación de ángulos de enlace realistas y potenciales torsionales.

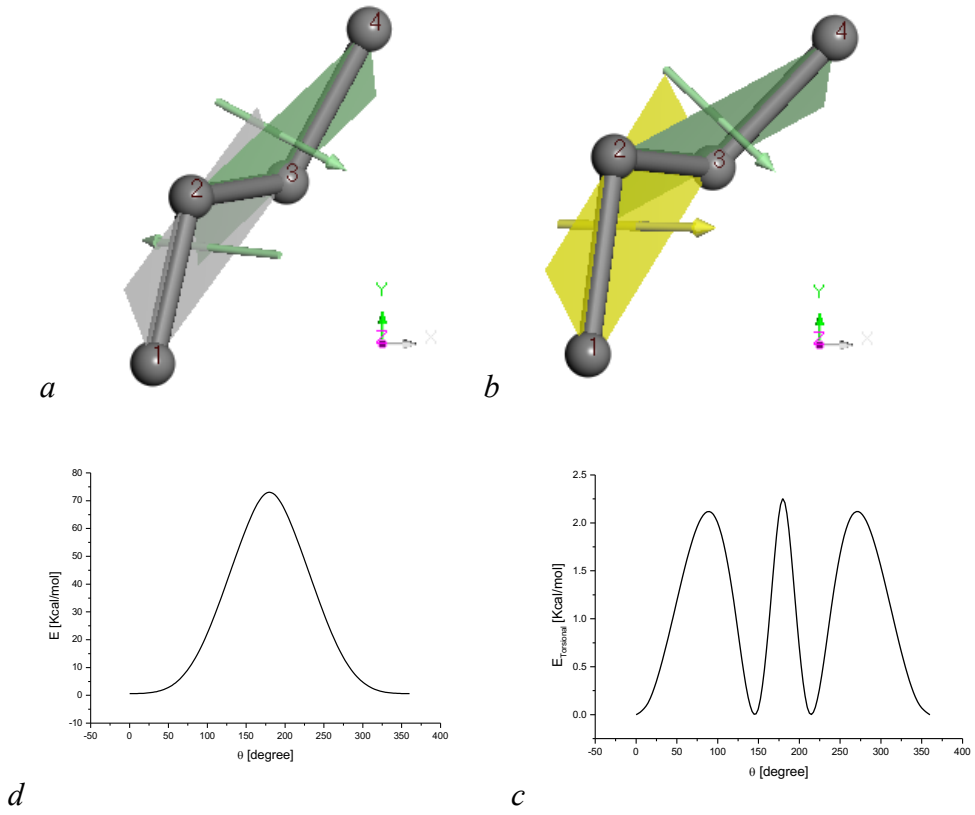
ii) Un fluido molecular de moléculas lineales, donde las interacciones electrostáticas permanentes son manejadas usando expansión multipolar.

iii) Un fluido de átomos con interacciones de tres-cuerpos modelados usando el potencial triple-dipolar Axilrod-Teller.

La torca entra en las ecuaciones rotacionales de movimiento de la misma manera que las fuerzas entran en las ecuaciones traslacionales; la naturaleza de la orientación en el espacio garantiza que las ecuaciones de movimiento reorientacional no serán tan simples como las ecuaciones traslacionales.

Para moléculas grandes puede que no sea razonable *fijar* todos los grados internos de libertad. En particular, el movimiento torsional alrededor de los enlaces, que conduce a interconversión conformacional, por ejemplo, alcanos, en general no pueden ser despreciados, ya que la energía cambia como se observa en la figura 39 (obtenida con Forcite de BIOVIA Materials Studio, 2017, usando Forcefield Universal). La simulación de n-butano hipotético solo de carbono proporciona un buen ejemplo de la manera en que estas características son incorporadas en una molécula simple. Butano puede representarse como una molécula de cuatro centros, con longitudes de enlace fijas y ángulos de flexión de enlace.

Figura 39. *a) Un modelo de butano. b) Mayor abertura entre los planos en el modelo de butano. c) Barrera de energía potencial. d) Barreras de energía potencial torsional*



Fuente: elaboración propia.

Una característica común es construir grupos de átomos como CH_3 y CH_2, que se condensan en *átomos unidos* esféricamente simétricos. Para butano, las interacciones entre dichos grupos se representan por el potencial de Lennard-Jones, con parámetros empíricos. En una simulación se toman fijas las longitudes de enlace C_1–C_2, C_2–C_3 y C_3–C_4 en la molécula de butano. Los ángulos θ (entre el eje C_2–C_3 y el enlace C_1–C_2) y θ′ (entre el eje C_2–C_3 y el enlace C_3–C_4), se fijan por constricción de las distancias C_1–C_3 y C_2–C_4 *i.e.* introduciendo *enlaces fantasmas*. De esta manera, queda un grado interno de libertad, definido por la rotación alrededor de C_2–C_3, medido por el ángulo ϕ; para cada molécula aparece un término extra en el hamiltoniano en la energía potencial $E_{torsional}(\phi)$ periódica en ϕ. Este potencial debería tener un mínimo (ó máximo) en un valor de ϕ que corresponde

al conformador *trans* de butano, y otro secundario en la conformación *gauche*. Esta aproximación se puede extender a moléculas flexibles más grandes.

Para manejar la dinámica de un sistema molecular ciertos grados de libertad (tales como longitudes de enlace) se han seleccionado arbitrariamente, mientras que otros quedan libres de evolucionar bajo la influencia de fuerzas intermoleculares e intramoleculares. Esta constricción dinámica usa un conjunto de multiplicadores indeterminados para representar magnitudes de las fuerzas dirigidas a lo largo de los enlaces que se requieren para mantener constantes las longitudes del enlace. Se trata de resolver las ecuaciones de movimiento para un paso de tiempo en ausencias de las fuerzas de constricción, y así determinar sus magnitudes y corregir las posiciones atómicas.

Considérese una molécula triatómica doblada tal como el agua, en la cual deseamos restringir a longitud fija dos de los enlaces, pero permitir el enlace restante, y que el ángulo inter-enlace varíe bajo la influencia del potencial intramolecular. Numerando 2 al átomo central (oxígeno), y 1 y 3 a los dos átomos exteriores (hidrógeno), las ecuaciones de movimiento tienen la forma

$$m_i \ddot{r}_1 = f_1 + g_1$$

$$m_i \ddot{r}_2 = f_2 + g_2$$

$$m_i \ddot{r}_3 = f_3 + g_3$$

Aquí $f_1, f_2, y\ f_3$ son las fuerzas debidas a interacciones intermoleculares y a aquellos efectos que se incluyen explícitamente en el potencial. Los términos restantes g_1 etcétera son las fuerzas de constricción: su papel es mantener constantes las longitudes de enlace deseadas, esto es para asegurar que las ecuaciones

$$\chi_{12} = r_{12}^2(t) - d_{12}^2 = 0$$

$$\chi_{23} = r_{23}^2(t) - d_{23}^2 = 0$$

(donde d_{12}^2 y d_{23}^2 son las longitudes de enlace, y $r_{12} = |r_1 - r_2|$) se satisfacen en todos los tiempos. Las ecuaciones lagrangianas de movimiento que se derivan de estas constricciones son las previas con:

$$g_a = \frac{1}{2}\lambda_{12}\nabla_{r_a}\chi_{13} + \frac{1}{2}\lambda_{23}\nabla_{r_a}\chi_{23}$$

y λ_{12} y λ_{23} son multiplicadores indeterminados. Los factores de (½) se introducen tal que los multiplicadores de Lagrange están de acuerdo con estas ecuaciones. Hasta ahora, no se han hecho aproximaciones, y en principio se pueden resolver las fuerzas de constricción. Sin embargo, dado que se deben resolver ecuaciones de movimiento aproximadamente, usando el método de diferencias finitas, las longitudes de enlace divergen constantemente de los valores deseados.

En lugar de esto, Ryckaert et al. (1978) sugirieron una aproximación en la cual las fuerzas de constricción se calculan para garantizar que las constricciones se satisfagan en cada paso de tiempo; lo que implica que las fuerzas de constricción mismas solo son correctas al mismo orden de exactitud que el algoritmo de integración. Entonces

$$m_a\ddot{r}_a = f_a + g_a \approx f_a + g_a^{(r)}$$

donde $g_a^{(r)}$ es una aproximación (de la cual posteriormente se dará su expresión) a las verdaderas fuerzas de constricción g_a que actúan en cada átomo a. Considerando la manera en que estas fuerzas entran en el algoritmo de Verlet se tiene que

$$r_a(t+\delta t) = r_a'(t+\delta t) + (\delta t^2/m_a)g_a^{(r)}(t)$$

donde $r_a'(t+\delta t)$ es la posición que alcanzaría en ausencia de cualquier constricción. Regresando al ejemplo del agua y reconociendo que las fuerzas de constricción deben dirigirse a los enlaces y deben conformar a la tercera ley de Newton, vemos que

$$g_1^{(r)} = \lambda_{12}r_{12}$$

$$g_2^{(r)} = \lambda_{23}r_{23} - \lambda_{12}r_{12}$$

$$g_3^{(r)} = -\lambda_{23}r_{23}$$

donde λ_{12} y λ_{23} son los multiplicadores indeterminados. Estos pueden calcularse si escribimos explícitamente la ecuación anterior de la posición, esto es

$$r_1(t+\delta t) = r_1'(t+\delta t) + (\delta t^2/m_1)\lambda_{12}r_{12}(t)$$

$$r_2(t+\delta t) = r_2'(t+\delta t) + (\delta t^2/m_2)\lambda_{23}r_{23}(t) - (\delta t^2/m_2)\lambda_{12}r_{12}(t)$$

$$r_3(t+\delta t) = r_3'(t+\delta t) - (\delta t^2/m_3)\lambda_{23}r_{23}(t)$$

Entonces

$$r_{12}(t+\delta t) = r'_{12}(t+\delta t) + \delta t^2\left(m_1^{-1} + m_2^{-1}\right)\lambda_{12} r_{12}(t) + \delta t^2\left(m_2^{-1} + m_3^{-1}\right)\lambda_{23} r_{23}(t)$$

$$r_{23}(t+\delta t) = r'_{23}(t+\delta t) - \delta t^2 m_2^{-1}\lambda_{12} r_{12}(t) + \delta t^2 (m_2^{-1} + m_3^{-1})\lambda_{23} r_{23}(t)$$

Tomando el módulo al cuadrado en ambos lados, y aplicando la constricción deseada: $|r_{12}(t+\delta t)|^2 = |r_{12}(t)|^2 = d_{12}^2$ y análogamente para r_{23}. El resultado es un par de ecuaciones cuadráticas en λ_{12} y λ_{23}, cuyos coeficientes son conocidos (dado que ya tenemos las posiciones r'_a *no constreñidas*) y que pueden resolverse para los multiplicadores indeterminados. Puesto que los términos de primer orden en λ_{12}, λ_{23} son proporcionales a δt^2, mientras que los términos de segundo orden son proporcionales a δt^4. En la práctica estas ecuaciones se resuelven iterativamente. Los términos cuadráticos se quitan y las ecuaciones lineales restantes se resuelven para λ_{12} y λ_{23}; estos valores se substituyen en los términos cuadráticos para dar nuevas ecuaciones lineales, que permiten estimaciones mejoradas de λ_{12} y λ_{23}, y así sucesivamente. Finalmente se usan estos valores en las ecuaciones explícitas de las posiciones.

La constricción (como opuesto a la longitud de enlace) ángulo de enlace no presenta dificultad fundamental y se puede manejar introduciendo constricciones de longitud adicional. Por ejemplo, el ángulo de enlace H-O-H en agua puede fijarse constriñendo la distancia H-H en adición a las longitudes de enlace O-H. Con estas constricciones las ecuaciones resultantes son:

$$r_1(t+\delta t) = r'_1(t+\delta t) + (\delta t^2/m_1)\lambda_{12} r_{12}(t) - (\delta t^2/m_1)\lambda_{31} r_{31}(t)$$

$$r_2(t+\delta t) = r'_2(t+\delta t) + (\delta t^2/m_2)\lambda_{23} r_{23}(t) - (\delta t^2/m_2)\lambda_{12} r_{12}(t)$$

$$r_3(t+\delta t) = r'_3(t+\delta t) + (\delta t^2/m_3)\lambda_{31} r_{31}(t) - (\delta t^2/m_3)\lambda_{23} r_{23}(t)$$

entonces

$$r_{12}(t+\delta t) = r'_{12}(t+\delta t) + \delta t^2 (m_1^{-1} + m_2^{-1})\lambda_{12} r_{12}(t) - \delta t^2 m_2^{-1}\lambda_{23} r_{23}(t) - \delta t^2 m_1^{-1}\lambda_{31} r_{31}(t)$$

$$r_{23}(t+\delta t) = r'_{23}(t+\delta t) - \delta t^2 m_3^{-1}\lambda_{31} r_{31}(t) + \delta t^2 (m_2^{-1} + m_3^{-1})\lambda_{23} r_{23}(t) - \delta t^2 m_2^{-1}\lambda_{12} r_{12}(t)$$

$$r_{31}(t+\delta t) = r'_{31}(t+\delta t) - \delta t^2 m_1^{-1}\lambda_{12} r_{12}(t) - \delta t^2 m_3^{-1}\lambda_{23} r_{23}(t) + \delta t^2 (m_3^{-1} + m_1^{-1})\lambda_{31} r_{31}(t)$$

Este proceso de *triangular* la molécula al introducir enlaces ficticios es directamente aplicado a sistemas más complicados. Las longitudes de enlace constreñidas aplicadas a las unidades de carbono en un modelo de butano, que deja solo un parámetro interno (el ángulo de torsión ϕ) libre a evolucionar bajo la influencia del potencial. La extensión a n-alcanos es discutida por Ryckaert et al. (1977) y una aplicación al caso de n-decano por Ryckaert y Bellemans (1978).

Para moléculas muy pequeñas, las ecuaciones de constricción (linealizadas) se pueden resolver por algebra directa. Para una molécula poliatómica más grande, con n_C constricciones, la solución de estas ecuaciones esencialmente requiere inversión de una matriz $n_C \times n_C$ en cada paso de tiempo. Esto puede consumir tiempo de cómputo en moléculas muy grandes tales como proteínas. Asumiendo que solo átomos y enlaces de vecinos cercanos se relacionan por ecuaciones de constricción, la matriz de constricción será escasa y se pueden aplicar técnicas de inversión especiales. Un procedimiento alternativo es ir cíclicamente una por una de las constricciones, ajustando las coordenadas tal que satisfagan cada una en turno. El procedimiento puede ser iterativo hasta que se satisfagan todas las iteraciones, dentro de una tolerancia dada (Allen y Tildesley, 1987; Ryckaert et al., 1977).

Se necesitan aproximaciones a las fuerzas de constricción verdaderas en las ecuaciones de movimiento para garantizar que las constricciones se satisfacen en todos los tiempos de paso. El algoritmo de velocidad de Verlet es un proceso de dos etapas, con cada etapa involucrando las fuerzas de constricción. En cada etapa se hace una aproximación a g tal que asegure que se satisfacen las constricciones. Las fuerzas de constricción entran en el algoritmo de Verlet como sigue. En la primera etapa tenemos

$$r_a(t+\delta t) = r'_a(t+\delta t) + \frac{1}{2}\left(\frac{\delta t^2}{m_a}\right) g_a^{(r)}(t)$$

$$v_a\left(t+\frac{1}{2}\delta t\right) = v'_a\left(t+\frac{1}{2}\delta t\right) + \frac{1}{2}\left(\frac{\delta t^2}{m_a}\right) g_a^{(r)}(t)$$

Las fuerzas de constricción $g_a^{(r)}$ para esta etapa están dirigidas a lo largo de los vectores de enlace $r_{ab}(t)$. Se determinan resolviendo las ecuaciones de la posición anteriores por

inversión matricial. Al mismo tiempo, las velocidades al tiempo $t + \frac{1}{2}\delta t$ se ajustan de acuerdo con la ecuación v_a'. En una segunda parte del algoritmo sigue la evaluación de las fuerzas $f_a(t + \delta t)$ que se usan para obtener $v_a'(t + \delta t)$. La segunda etapa es:

$$v_a(t + \delta t) = v_a'(t + \delta t) + \frac{1}{2}\left(\frac{\delta t^2}{m_a}\right) g_a^{(v)}(t + \delta t)$$

Estas fuerzas de constricción $g_a^{(v)}(t + \delta t)$ se dirigen a lo largo de los enlaces $r_{ab}(t + \delta t)$, y se eligen tal que las velocidades satisfacen las constricciones exactamente al tiempo $t + \delta t$. Nótese que en el siguiente paso de integración se usará una aproximación diferente a estas mismas fuerzas de constricción, a saber, $g_a^{(r)}(t + \delta t)$.

Cuando se introduce no-rigidez es conveniente usar coordenadas generalizadas. Para moléculas flexibles existe una elección amplia correspondiente a donde aplicar constricciones, y sin restar importancia a las longitudes de enlace que constriñen, los ángulos de enlace se dejan libres a evolucionar bajo la influencia de términos apropiados en la energía potencial.

Moléculas no-lineales

La orientación de un cuerpo rígido especifica la relación entre un sistema de ejes fijos en el espacio y uno fijo con respecto al cuerpo, usualmente el sistema *principal* es el que está fijo-al-cuerpo, en el cual el tensor de inercia es diagonal (Evans, 1977). Cualquier vector unitario $\hat{e}$ se puede expresar en términos de componentes en el sistema fijo-en-el-cuerpo o fijo-en-el-espacio: se usa la notación de vectores unitarios $\hat{e}^b$ y $\hat{e}^s$, respectivamente. Estas componentes están relacionadas por la matriz de rotación A

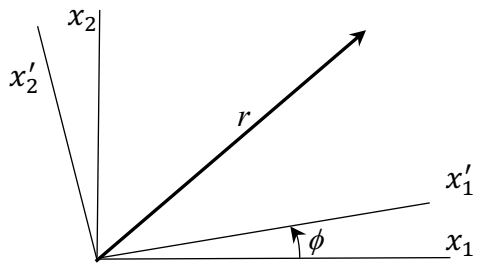
$$\hat{e}^b = A \cdot \hat{e}^s$$

Que en nuestro caso corresponden a

$$r' = Ar \Rightarrow \begin{cases} x' = a_{11}x + a_{12}y + a_{13}z \\ y' = a_{21}x + a_{22}y + a_{23}z \\ z' = a_{31}x + a_{32}y + a_{33}z \end{cases} \quad \text{redefinimos} \begin{cases} x = x_1 \\ y = y_1 \\ z = z_1 \end{cases}$$

En un plano se tiene (figura 40):

Figura 40. *Transformación de coordenadas*



Fuente: elaboración propia.

$$x_1' = x_1 cos\phi + x_2 sen\phi$$

$$x_2' = -x_1 sen\phi + x_2 cos\phi$$

$$\therefore \qquad x' = Ax$$

donde $A = \begin{pmatrix} cos\phi & sen\phi \\ -sen\phi & cos\phi \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $x' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$

Consideremos lo que ocurre cuando dos transformaciones sucesivas corresponden a dos desplazamientos sucesivos del cuerpo rígido. Denotemos la primera transformación de $r\ a\ r'$ por B:

$$x_k' = b_{kj} x_j$$

Y la transformación sucesiva de r' a un tercer conjunto de coordenadas r'' por A es:

$$x_i'' = a_{ik} x_k'$$

Combinando estas dos ecuaciones se tiene

$$x_i'' = a_{ik} b_{kj} x_j$$

o bien $$x_i'' = c_{ij} x_j'$$

donde $$c_{ij} = a_{ik} b_{kj} \ \Rightarrow \qquad C = AB$$

La aplicación sucesiva de dos transformaciones ortogonales A, B es equivalente a una tercera transformación lineal C, también ortogonal. En la figura 41 se indican tres transformaciones ortogonales sucesivas

$$x_i' = d_{ij} x_j$$

$$x_j'' = c_{ki} x_i'$$

$$x_k''' = b_{ik} x_j''$$

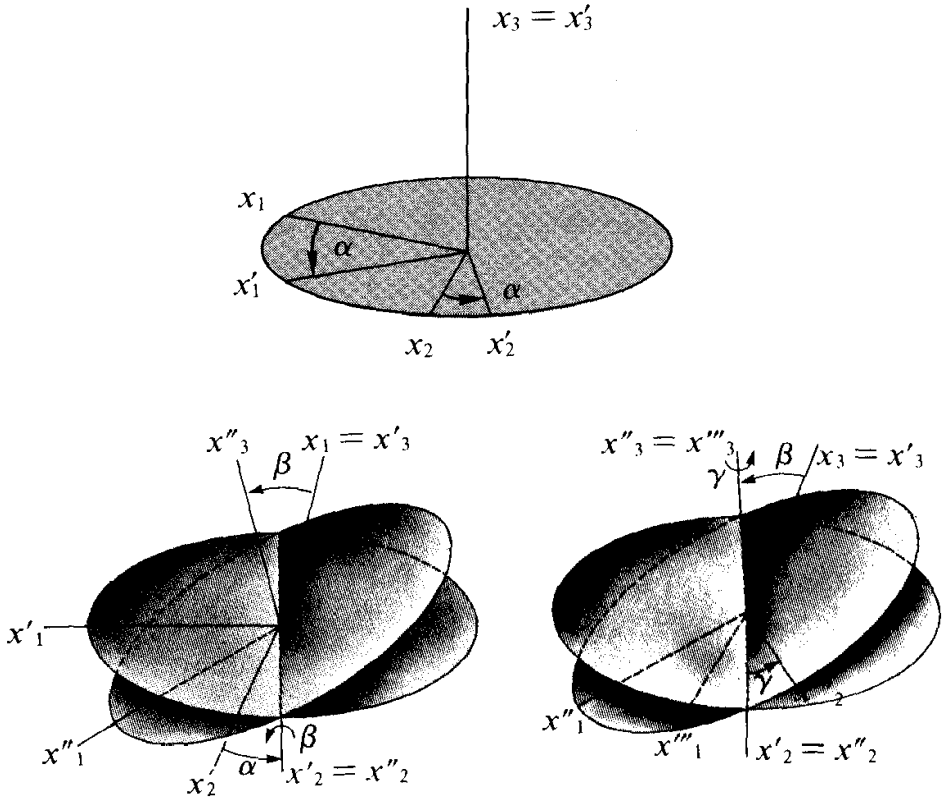
$$x_k''' = b_{ik} c_{ki} d_{ij} x_j \qquad \rightarrow \qquad x_k''' = b_{ik} e_{kj} x_j \qquad \rightarrow \qquad x_k''' = a_{ij} x_j$$

$$a_{ij} = b_{ik} c_{ki} d_{ij} \qquad \rightarrow \qquad A = BCD$$

La transformación *D* es una rotación de un ángulo α (o ϕ) alrededor de z ($o\ x_3$), y tiene una matriz:

$$D = \begin{pmatrix} cos\phi & sen\phi & 0 \\ -sen\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Figura 41. *Cosenos directores de vectores de ejes fijos-en-el-cuerpo en el sistema de referencia del espacio-fijo*



Fuente: elaboración propia.

La transformación *C* corresponde a una rotación de un ángulo β (o θ) alrededor de *y* ($o\ x_2$), y su matriz es

$$C = \begin{pmatrix} cos\theta & 0 & -sen\theta \\ 0 & 1 & 0 \\ sen\theta & 0 & cos\theta \end{pmatrix}$$

y finalmente *B* es otra rotación a un ángulo γ (o ψ) alrededor del nuevo z (o x'''_3), con matriz:

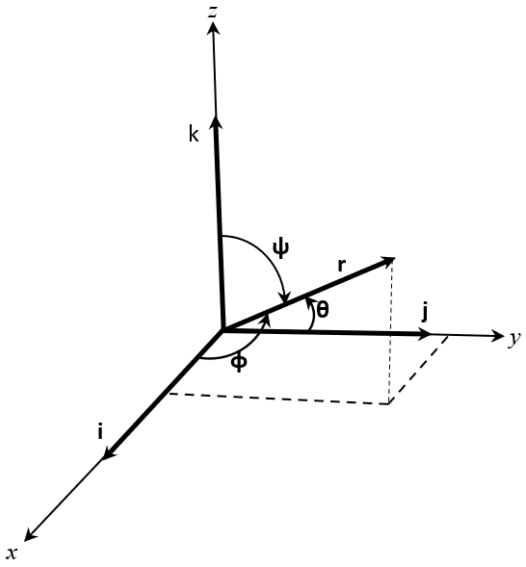
$$B = \begin{pmatrix} cos\psi & sen\psi & 0 \\ -sen\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

El producto de las tres matrices en la forma A=DBC da

$$A = \begin{pmatrix} -sen\phi sen\psi + cos\phi cos\theta cos\psi & sen\phi cos\psi + cos\phi cos\theta sen\psi & -cos\phi sen\theta \\ -cos\phi sen\psi - sen\phi cos\theta cos\psi & cos\phi cos\psi - sen\phi cos\theta sen\psi & sen\phi sen\psi \\ sen\theta cos\psi & sen\theta sen\psi & cos\theta \end{pmatrix}$$

Las nueve componentes de la matriz de rotación son los cosenos directores de los vectores de ejes fijos-en-el-cuerpo en el sistema de referencia de espacio-fijo, y definen completamente la orientación molecular. De hecho, se necesitan tres cantidades independientes (coordenadas generalizadas) para definir A, que pueden ser los ángulos de Euler: $\phi\theta\psi$. Véase la figura 42 sustituyendo los ángulos $\alpha\beta\gamma$ por $\phi\theta\psi$.

Figura 42. *Cosenos directores*



Fuente: elaboración propia.

Precisión

Primero se debe checar que las leyes de conservación son apropiadamente obedecidas, y en particular que la energía sea constante. Para un sistema de Lennard-Jones, fluctuaciones del orden de 1 en 10^4 generalmente se consideran aceptables. Las fluctuaciones en la energía se pueden reducir decreciendo el tiempo de paso. Si se usa uno de los algoritmos de Verlet, se deberían tomar varias corridas cortas comenzando desde la misma configuración inicial, y cubriendo el mismo tiempo total t_{run}: cada corrida debe emplear un tiempo de paso δt

diferente, y consiste en un número de pasos $\tau_{run} = \frac{t_{run}}{\delta t}$. Las fluctuaciones RMS de energía para cada corrida deben calcularse. Si el programa funciona correctamente, y otras fuentes de fluctuaciones de energía (como truncamiento del potencial) se han eliminado, entonces el algoritmo de Verlet debe dar fluctuaciones RMS de energía que son proporcionales a δt^2.

Nótese que la mayoría de los problemas son *mecánicos* en vez de *termodinámicos*, *i.e.* están asociados con la solución correcta de la ecuación de movimiento. Una cuestión para pensarse es si la distancia de corte del potencial es más pequeña que la longitud de la caja en un sistema periódico.

Dinámica molecular de esferas duras

Las colisiones entre partículas conducen a diversas situaciones, entre lo más simple está abordar este caso inicialmente por medio de los principios de conservación de la energía y del momento, en un campo conservativo. Para los cuales está claramente demostrado que los cambios en momento de dos cuerpos son iguales y opuestos en interacciones elásticas

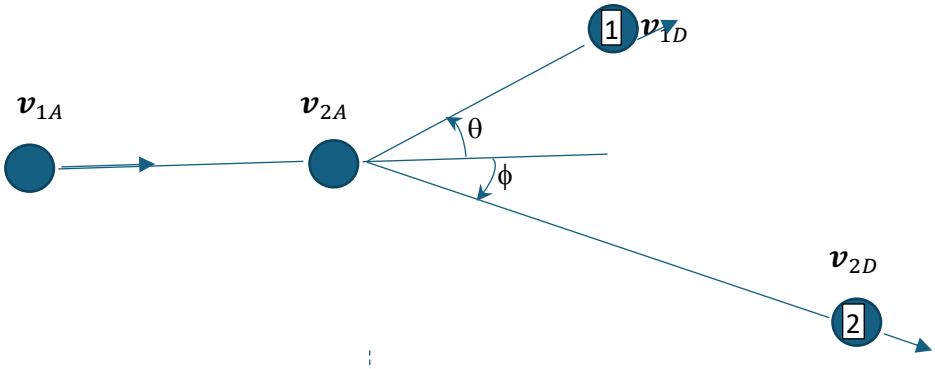
$$-\Delta \boldsymbol{p}_A = \Delta \boldsymbol{p}_D$$

Así mismo, los cambios en sus energías cinéticas son iguales y opuestos

$$-\Delta E_{K_A} = \Delta E_{K_D}$$

Se necesitan el momento y la energía de cada partícula antes y después de la colisión. Esto se ilustra con el ejemplo de una colisión elástica entre dos canicas iguales y de la misma masa en un plano con la canica 1 dirigiéndose a la canica 2 que está fija.

Figura 43. *Colisión de dos canicas de la misma masa, en un plano*



Fuente: elaboración propia.

La canica 2 está en reposo antes de la colisión, por lo cual $\boldsymbol{v}_{2A} = 0$. Después de la colisión la canica 1 se mueve con un ángulo θ respecto a la dirección de la colisión, y la canica 2 con un ángulo ϕ como se ve en la figura 43. Dado que las canicas tienen masas iguales, y debido al principio de conservación del momento ya mencionado aquí arriba, se obtienen las ecuaciones

$$v_{1A} = v_{1D}cos\theta + v_{2D}cos\phi$$

$$0 = v_{1D}sen\theta - v_{2D}sen\phi$$

La primera ecuación se toma en la dirección horizontal, la segunda en la dirección vertical; y dadas las masas y la velocidad inicial se requiere de una de las cantidades restantes: $v_{1D}, v_{2D}, \theta\ y\ \phi$. Entonces, una tercera ecuación se obtiene por el principio de conservación de la energía:

$$v_{1A}^2 = v_{1D}^2 + v_{2D}^2.$$

A partir de estas tres ecuaciones, en este caso, por ejemplo, se puede deducir la siguiente relación

$$sen^2(\phi + \theta) = sen^2\phi + sen^2\theta$$

$$Si\ \phi + \theta = \frac{\pi}{2}, \qquad entonces \qquad 1 = sen^2\phi + sen^2\theta$$

y por relaciones trigonométricas se concluye que $sen^2\phi = cos^2\theta$

Con este ejemplo se muestra que en mecánica clásica se pueden evidenciar las colisiones elásticas en campos conservativos.

Ahora bien, el desarrollo de programación computacional de potenciales de esferas duras progresa en base a colisión-por-colisión, computando la dinámica de colisiones y luego investigando la siguiente colisión, en general se tiene que:

a) Localizar la siguiente colisión.
b) Mover todas las partículas hacia adelante hasta que ocurra la colisión.
c) Implementar dinámica de colisión para el par en colisión.
d) Calcular cualquier propiedad de interés, fácil de promediar, antes de regresar a (a).

Considerando dos esferas duras *i* y *j*, de diámetro σ, cuyas posiciones en el tiempo son $\boldsymbol{r}_i$ y $\boldsymbol{r}_j$, y cuyas velocidades son $\mathbf{v}_i$ y $\mathbf{v}_j$, respectivamente. Si estas partículas colisionan elásticamente al tiempo $t + t_{ij}$, entonces se satisface que el diámetro de la esfera esta dado por

$$\left|\boldsymbol{r}_{ij}\left(t + t_{ij}\right)\right| = \left|\boldsymbol{r}_{ij} + \mathbf{v}_{ij}t_{ij}\right| = \sigma$$

donde $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$ y $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$, nótese que $\boldsymbol{r}_{ij}$ va del centro de la partícula *j* al centro de la partícula *i*, como en la figura 44. Si $b_{ij} = \boldsymbol{r}_{ij} \cdot \mathbf{v}_{ij}$, entonces esta magnitud es

$$v_{ij}^2 t_{ij}^2 + 2b_{ij}t_{ij} + r_{ij}^2 - \sigma^2 = 0$$

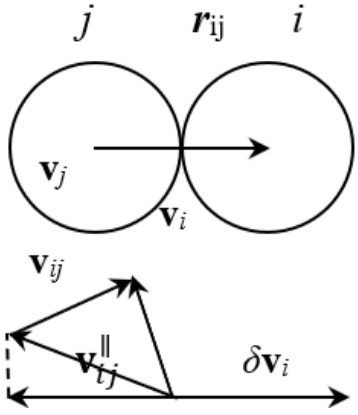
La cual es una ecuación cuadrática en t_{ij} con solución

$$t_{ij} = \frac{-b_{ij} \pm \left(b_{ij}^2 - v_{ij}^2\left(r_{ij}^2 - \sigma^2\right)\right)^{1/2}}{v_{ij}^2}$$

Si $b_{ij} > 0$, las partículas se alejan entre sí y no colisionan. Si $b_{ij} < 0$ aun podría ser verdadero que $b_{ij}^2 - v_{ij}^2\left(r_{ij}^2 - \sigma^2\right) < 0$ en caso de que las raíces sean complejas, y no hay colisión. Por otra parte, la más pequeña de las dos raíces corresponde a impacto.

Los cambios de velocidades del par que chocan dependen de que la energía y el momento lineal se conserven, y que el impulso actúe a lo largo de la línea de centros como se muestra en la figura 44.

Figura 44. *Colisión suave de esferas-duras*



Nota. La construcción del vector da los cambios de velocidades para cada partícula. Se toman coplanares los vectores (Allen y Tildesley, 1987).

Fuente: elaboración propia.

Usando conservación del momento lineal total y de la energía (cinética), y tomando masas iguales, el cambio de velocidades $\delta\mathbf{v}_i$, tal que

$$\mathbf{v}_i(después) = \mathbf{v}_i(antes) + \delta\mathbf{v}_i$$

$$\mathbf{v}_j(después) = \mathbf{v}_j(antes) + \delta\mathbf{v}_j$$

está dado por

$$\delta\mathbf{v}_i = -\left(\frac{b_{ij}}{\sigma^2}\right)\boldsymbol{r}_{ij} = -\mathbf{v}_{ij}^{\parallel}$$

con b_{ij} ahora evaluado en el momento de impacto. Entonces, $\delta\mathbf{v}_i$ simplemente es el negativo de la proyección de $\mathbf{v}_{ij}$ a lo largo de $\boldsymbol{r}_{ij}$, la cual se denota por $\mathbf{v}_{ij}^{\parallel}$.

Ahora, para cada par hay dos distancias en las cuales ocurren *colisiones*. Las colisiones en el interior de la esfera obedecen dinámica de esfera-dura; en la frontera exterior, el cambio en el momento se determina por la ley de conservación usual. Para moléculas aproximándose entre sí, la energía potencial decae al cruzar la frontera, y la energía cinética muestra un incremento correspondiente. Si las moléculas se separan se tienen dos posibilidades: 1) Si la energía cinética total es suficiente, las moléculas cruzan la frontera con una pérdida de energía cinética para compensar el alza en energía potencial. 2) Si la energía cinética es suficiente, ocurre reflexión en la frontera exterior y las partículas permanecen *pegadas*.

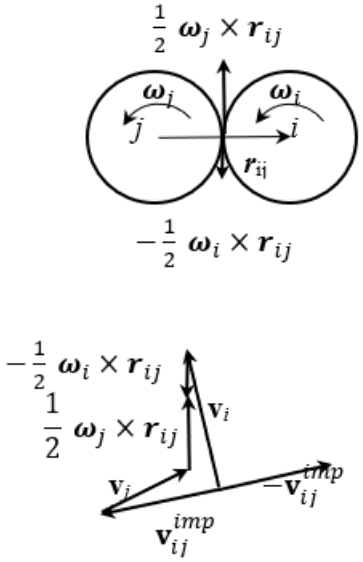
Otra modificación al modelo de esferas-duras que preserva simetría esférica es la rugosidad. Las esferas rugosas difieren de esferas duras lisas en su dinámica de colisión: la dinámica de vuelos libres entre colisiones son idénticas. Esferas rugosas se caracterizan por un diámetro σ, una masa m, y un momento de inercia I, o, alternativamente, un parámetro de energía cinética $\mathcal{K} = 4I/m\sigma^2$. Tienen velocidades traslacionales $\mathbf{v}$ y velocidades angulares $\boldsymbol{\omega}$ o *spin*. Las colisiones dinámicas de esferas rugosas están sujetas a las leyes de conservación usuales: una colisión entre dos partículas preserva energía total (rotacional + transnacional), momento lineal total, y momento angular total definido por

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S} = \sum m\boldsymbol{r}_i \times \mathbf{v}_i + \sum \boldsymbol{I}\boldsymbol{\omega}_i$$

La diferencia entre esferas duras rugosas y suaves se ve en la figura 45. Considere los dos puntos en las superficies de las esferas que se juntan en el momento del impacto. El vector de velocidad relativa de estos puntos es

$$\mathbf{v}_{ij}^{imp} = \mathbf{v}_i^{imp} - \mathbf{v}_j^{imp} = (\mathbf{v}_i - \mathbf{v}_j) - \frac{1}{2}(\boldsymbol{\omega}_i + \boldsymbol{\omega}_j) \times \boldsymbol{r}_{ij}$$

Figura 45. *Una colisión de esfera dura rugosa*



Nota. Se muestra la inversa del vector de velocidad relativo de los puntos de impacto. Para propósitos ilustrativos, todos los vectores de velocidad angular que apuntan hacia arriba normal al plano (Allen y Tildesley, 1987).

Fuente: elaboración propia.

Se puede resolver este vector en componentes paralelas a la línea de los centros del par en colisión, y perpendicular a esta línea

$$\mathbf{v}_{ij}^{imp} = \mathbf{v}_{ij}^{imp\parallel} + \mathbf{v}_{ij}^{imp\perp}$$

debido a que las esferas son duras, la componente paralela de $\mathbf{v}_{ij}^{imp}$ es inversa en impacto. Para esferas duras suaves, la parte perpendicular no se altera, y entonces no hay cambio en las velocidades angulares. Para esferas duras rugosas ambas partes del vector de velocidad relativa $\mathbf{v}_{ij}^{imp}$ son inversas en impacto. En consecuencia

$$\delta \mathbf{v}_{ij}^{imp} = \mathbf{v}_{ij}^{imp}(después) - \mathbf{v}_{ij}^{imp}(antes) = -2\mathbf{v}_{ij}^{imp}(antes)$$

Entonces las leyes de conservación conducen a una expresión para el impulso

$$\delta \boldsymbol{p_i} = \frac{1}{2} m \left(\delta \mathbf{v}_{ij}^{imp\parallel} + \frac{\kappa}{1+\kappa} \delta \mathbf{v}_{ij}^{imp\perp} \right)$$

En términos de los cuales los cambios en velocidades angulares se convierten en:

$$m\mathbf{v}_i(después) = m\mathbf{v}_i(antes) + \delta \boldsymbol{p_i}$$

$$m\mathbf{v}_j(después) = m\mathbf{v}_j(antes) - \delta \boldsymbol{p_i}$$

$$I\boldsymbol{\omega}_i(después) = I\boldsymbol{\omega}_i(antes) - \frac{1}{2}\boldsymbol{r}_{ij} \times \delta \boldsymbol{p_i}$$

$$I\boldsymbol{\omega}_j(después) = I\boldsymbol{\omega}_j(antes) - \frac{1}{2}\boldsymbol{r}_{ij} \times \delta \boldsymbol{p_i}$$

Para un sistema de esferas duras, aunque $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ se conserva en encuentros moleculares, el momento angular intrínseco $\boldsymbol{S}$ o *spin* no se conserva. En este caso cada colisión crea (o destruye) *spin*, y el cambio está dividido igualmente entre los demás. Consecuentemente $\boldsymbol{L}$ tampoco se conserva separadamente. Nótese que las condiciones de frontera periódicas destruyen las leyes de conservación para $\boldsymbol{J}$.

Capítulo 4. Modelación y simulación con método de Montecarlo

El método de Montecarlo es un método numérico que permite resolver problemas matemáticos mediante la simulación de variables aleatorias. Para tener una idea más clara usemos dos métodos (uno es triángulos, y el otro es Montecarlo) para obtener el valor de $\pi = \frac{perímetro}{diámetro}$ de una circunferencia, el cual tiene un número infinito de decimales, pero en seis dígitos de aproximación es 3.14159.

Primero insertamos triángulos en polígonos regulares para calcular el área del polígono en turno, es decir, se aproxima su valor utilizando triángulos dentro de un polígono al que sabemos que conforme crece el número de lados tiende al círculo circunscrito, ya que que el área de un círculo de radio uno es justamente π (A=πr^2, r=1). Empezamos calculando el área de un triángulo equilátero, el radio *r* de un círculo circunscrito es menor que la longitud de uno de sus lados. El área de un triángulo es la mitad de la base *b* por la altura *h*:

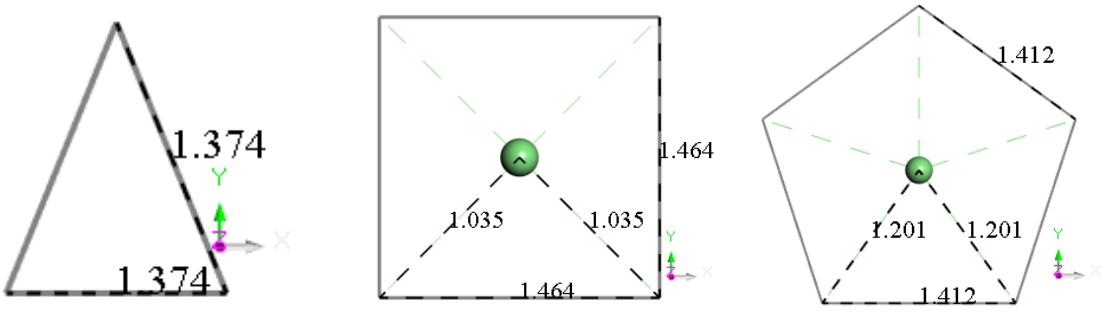
$$A_t = \frac{1}{2}bh$$

El área aproximada de un círculo de radio igual a uno es

$$A \cong nA_t$$

Donde n es el número de triángulos dentro del círculo. En la figura 46 se tienen los tres polígonos más pequeños. Las unidades que aparecen están en *angstroms*, y demuestran que los lados son iguales.

Figura 46. *Triángulo, cuadrado y pentágono de lados iguales*



Nota. En el cuadrado y en el pentágono la distancia de un vértice al centroide es el radio, el cual es menor que la longitud de un lado.

Fuente: elaboración propia.

La altura en el triángulo equilátero la calculamos usando el teorema de Pitágoras $c^2 = a^2 + b^2$. Para esto tomemos *x* como un lado del triángulo, entonces

$$x^2 = \left(\frac{1}{2}x\right)^2 + h^2$$

$$h^2 = \left(1 - \left(\frac{1}{2}\right)^2\right)x^2$$

$$h = \frac{\sqrt{3}}{2}x$$

Sustituyendo en el área de un triángulo de base *x*, se tiene que

$$A_t = \frac{1}{2}x\left(\frac{\sqrt{3}}{2}x\right) = \frac{\sqrt{3}}{4}x^2$$

Como $r < x$, y suponiendo $x = 1$, obtenemos el área de un círculo de $r < 1$, que tomamos como una primera aproximación al valor de π

$$A \cong A_t = \frac{\sqrt{3}}{4} = 0.433$$

Para el cuadrado, como se ve en la figura 46, tenemos cuatro triángulos inscritos. Cada triángulo tiene $h = \frac{1}{2}x$, y base *x*, por tanto, el área de cada triángulo es

$$A_t = \frac{1}{2}x\left(\frac{1}{2}x\right) = \frac{1}{4}x^2$$

$$\therefore \quad A \cong 4A_t = 4\left(\frac{1}{4}x^2\right) = x^2$$

Como $r < x$, y suponiendo $x = 1$, obtenemos el área de un círculo con $r < 1$. Como una segunda aproximación al valor de π tenemos

$$A \cong 1$$

Para el pentágono suponemos $x = 1$, dado que $r < x$, y usando los valores del pentágono en la figura 46, se tiene que $r = \frac{1.201}{1.412}x$, en consecuencia, por el teorema de Pitágoras

$$\left(\frac{1.201}{1.412}x\right)^2 = \left(\frac{1}{2}x\right)^2 + h^2$$

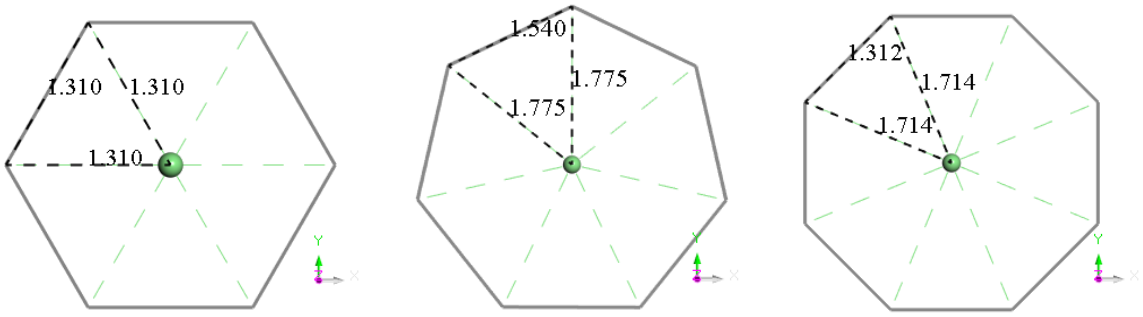
$$h^2 = \left(\left(\frac{1.201}{1.412}\right)^2 - \left(\frac{1}{2}\right)^2\right)x^2$$

$$h = 0.688x$$

$$A_t = \frac{1}{2}x(0.688x) = 0.344x^2$$

$$\therefore \quad A \cong 5A_t = 1.72$$

Figura 47. *Hexágono, heptágono y octágono de lados iguales*



Nota. La distancia de un vértice al centroide es el radio de un círculo circunscrito.

Fuente: elaboración propia.

Para el hexágono los seis triángulos tienen lados iguales; por lo tanto, por el teorema de Pitágoras

$$x^2 = \left(\frac{1}{2}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\right)^2\right)x^2$$

$$h = \frac{\sqrt{3}}{2}x$$

$$A_t = \frac{1}{2}x\left(\frac{\sqrt{3}}{2}x\right) = \frac{\sqrt{3}}{4}x^2$$

$$\therefore \quad A \cong 6A_t = \frac{3\sqrt{3}}{2} = 2.598$$

Para el heptágono suponemos $r \geq x, r = 1$. Usando los valores del heptágono en la figura 47, se tiene que $x = \frac{1.540}{1.775}r$, y tomando el límite cuando $r = x$, por el teorema de Pitagoras se tiene que

$$x^2 = \left(\frac{1}{2}\frac{1.540}{1.775}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\frac{1.540}{1.775}\right)^2\right)x^2$$

$$h^2 = 0.8118x^2$$

$$h = 0.901007833x$$

$$A_t = \frac{1}{2}\frac{1.540}{1.775}x(0.901007833x) = 0.390859736x^2$$

$$\therefore \quad A \cong 7A_t = 2.736$$

Para el octágono suponemos $r \geq x, r = 1$. Usando los valores del octágono en la figura 47 se tiene que $x = \frac{1.312}{1.714}r$, y tomando el límite cuando $r = x$, por el teorema de Pitagoras se tiene que

$$x^2 = \left(\frac{1}{2}\frac{1.312}{1.714}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\frac{1.312}{1.714}\right)^2\right)x^2$$

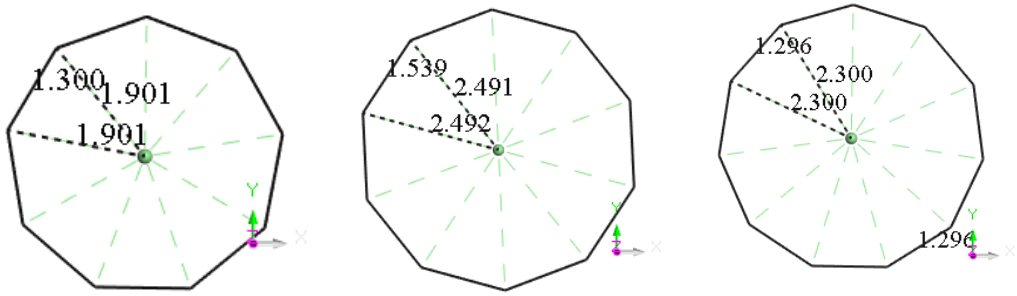
$$h^2 = 0.853517398x^2$$

$$h = 0.923860053x$$

$$A_t = \frac{1}{2}\frac{1.312}{1.714}x(0.923860053x) = 0.353589378x^2$$

$$\therefore \quad A \cong 8A_t = 2.82871503$$

Figura 48. *Nonágono, decágono y endecágono de lados iguales*



Nota. La distancia de un vértice al centro es el radio de un círculo circunscrito.

Fuente: elaboración propia.

Para el nonágono suponemos $r \geq x, r = 1$. Usando los valores del nonágono en la figura 48, se tiene que $x = \frac{1.300}{1.901}r$, y tomando el límite cuando $r = x$, por el teorema de Pitagoras se tiene que

$$x^2 = \left(\frac{1}{2}\frac{1.300}{1.901}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\,\frac{1.300}{1.901}\right)^2\right)x^2$$

$$h^2 = 0.883087087x^2$$

$$h = 0.939727134x$$

$$A_t = \frac{1}{2}\,\frac{1.300}{1.901}x(0.939727134x) = 0.321316484x^2$$

$$\therefore \quad A \cong 9A_t = 2.891848363$$

Para el decágono suponemos $r \geq x$, $r = 1$. Al usar los valores del decágono en la figura 48 se tiene que $x = \frac{1.539}{2.491}r$, y tomar el límite cuando $r = x$, por el teorema de Pitagoras se tiene que

$$x^2 = \left(\frac{1}{2}\,\frac{1.539}{2.491}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\,\frac{1.539}{2.491}\right)^2\right)x^2$$

$$h^2 = 0.904573324x^2$$

$$h = 0.951090587x$$

$$A_t = \frac{1}{2}\,\frac{1.539}{2.491}x(0.951090587x) = 0.293803378x^2$$

$$\therefore \quad A \cong 10A_t = 2.938033781$$

Para el endecágono suponemos $r \geq x$, $r = 1$. Usando los valores del endecágono en la figura 48 se tiene que $x = \frac{1.296}{2.3}r$, y tomando el límite cuando $r = x$, por el teorema de Pitagoras se tiene que

$$x^2 = \left(\frac{1}{2}\,\frac{1.296}{2.3}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\,\frac{1.296}{2.3}\right)^2\right)x^2$$

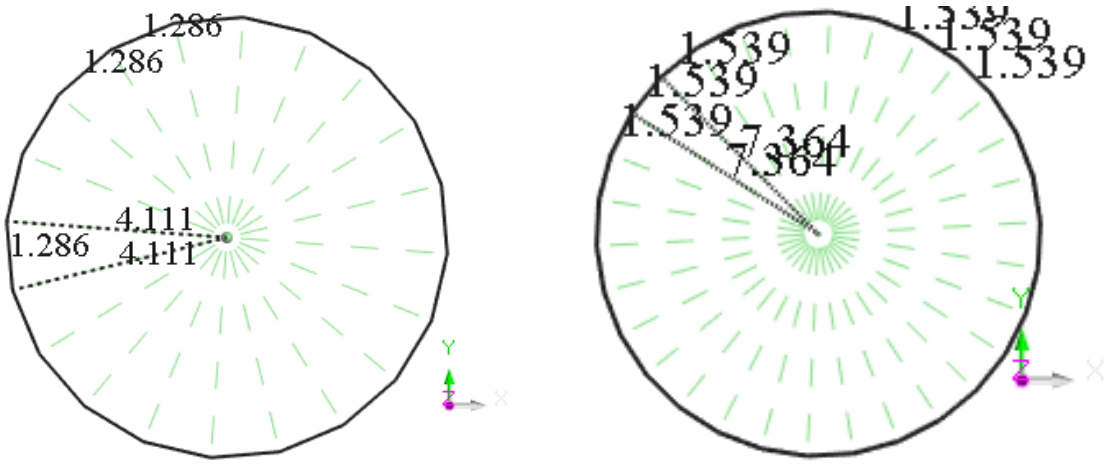
$$h^2 = 0.920623062x^2$$

$$h = 0.959491043x$$

$$A_t = \frac{1}{2}\,\frac{1.296}{2.3}x(0.959491043x) = 0.270326172x^2$$

$$\therefore \quad A \cong 11A_t = 2.973597895$$

Figura 49. *Isodecágono y triacontágono*



Fuente: elaboración propia.

Para el isodecágono suponemos $r \geq x$, $r = 1$. Usando los valores del isodecágono en la figura 49 se tiene que $x = \frac{1.286}{4.111} r$, y tomando el límite cuando $r = x$, por el teorema de Pitagoras se tiene que

$$x^2 = \left(\frac{1}{2}\frac{1.286}{4.111}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\frac{1.286}{4.111}\right)^2\right)x^2$$

$$h^2 = 0.975536026x^2$$

$$h = 0.987692273x$$

$$A_t = \frac{1}{2}\frac{1.286}{4.111}x(0.987692273x) = 0.154484585x^2$$

$$\therefore \quad A \cong 20A_t = 3.089691714$$

Para el triacontágono suponemos $r \geq x$, $r = 1$. Usando los valores del triacontágono en la figura 49 se tiene que $x = \frac{1.539}{7.364} r$, y tomando el límite cuando $r = x$, por el teorema de Pitagoras se tiene que

$$x^2 = \left(\frac{1}{2}\frac{1.539}{7.364}x\right)^2 + h^2$$

$$h^2 = \left(1^2 - \left(\frac{1}{2}\frac{1.539}{7.364}\right)^2\right)x^2$$

$$h^2 = 0.989080828x^2$$

$$h = 0.994525428x$$

$$A_t = \frac{1}{2}\,\frac{1.539}{7.364}x(0.994525428x) = 0.103922775x^2$$

$$\therefore \qquad A \cong 30A_t = 3.11768326$$

La secuencia para calcular π usando triángulos en el interior del círculo es la siguiente

$$A_t = triángulo, cuadrado, pentágono, hexágono, heptágono, octágono, \ldots,$$

$$A_t = 1A_t, 4A_t, 5A_t, 6A_t, 7A_t, 8A_t, 9A_t, 10A_t, 11A_t, \ldots, 20A_t, \ldots, 30A_t, \ldots$$

$$A_t = \frac{\sqrt{3}}{4}, 1, \frac{43}{25}, \frac{3\sqrt{3}}{2}, 2\frac{92}{125}, 2.8287, 2.8918, 2.9380, 2.9736, \ldots, 3.0897, \ldots, 3.1177, \ldots$$

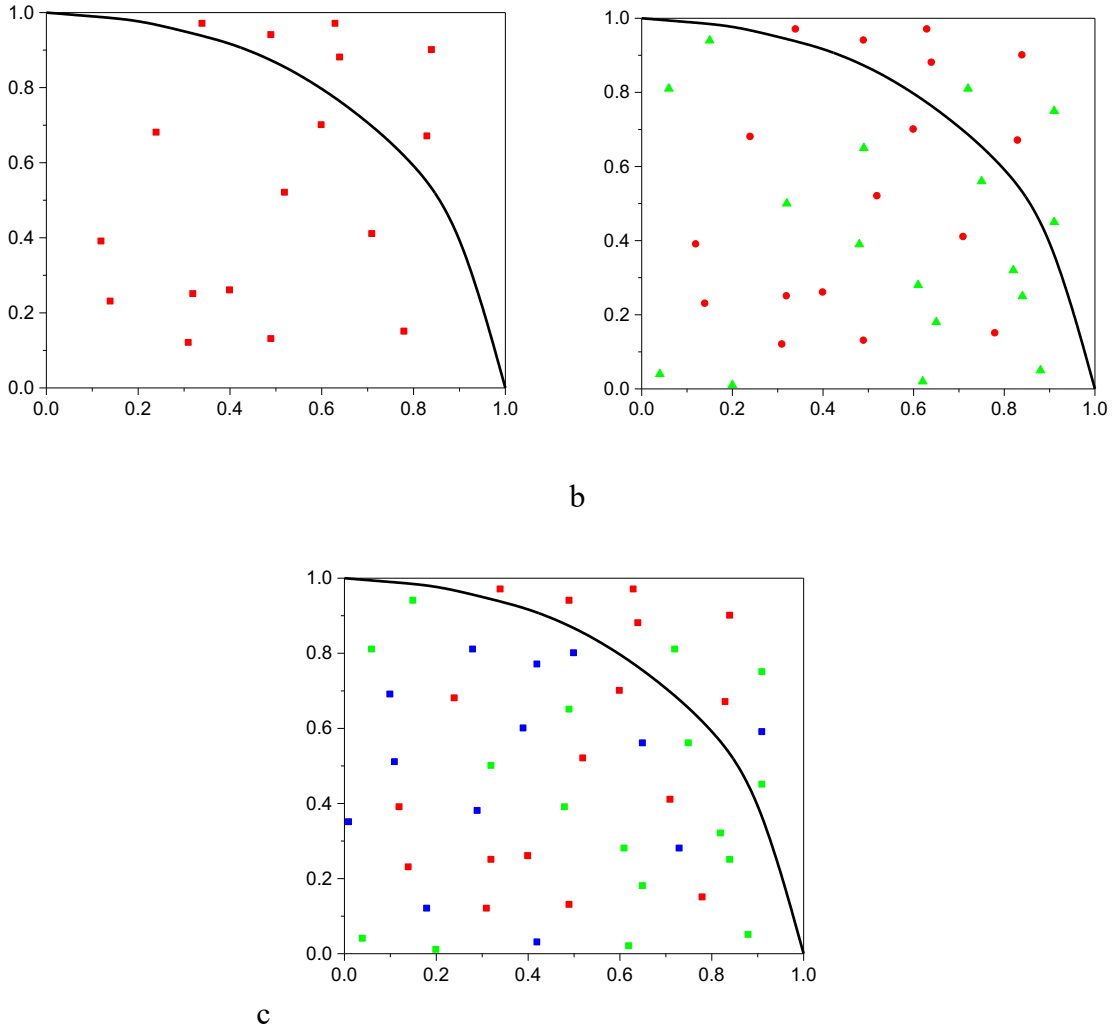
Con este método tan extenuante solo hemos conseguido dos dígitos para el valor de π con un polígono regular de 30 lados.

Segundo: método de Montecarlo. Apliquemos este método que considera aproximar el valor π usando una distribución uniforme de puntos. Esto es, en un cuadrante de un plano imagen tiene 17, otra tiene 34 y la otra tiene 47 puntos distribuidos uniformemente (figuras 50a, 50b, 50c, respectivamente). Al dividir el número de puntos en el interior del cuadrante de círculo entre el número total de puntos dentro del cuadrado de lado 1, se tiene que

$$\frac{11}{17} = 0.647$$

multiplicando este resultado por cuatro (4×0.647=2.588) completamos el círculo, entonces la primera aproximación a π resulta ser igual a 2.588. Hacemos lo mismo para la imagen que tiene 34 puntos y tenemos

Figura 50. *Obtención de π con Montecarlo*



a b

c

Fuente: elaboración propia.

$$\frac{25}{34} = 0.735$$

y ahora la aproximación a π es 2.94. Finalmente, para la imagen que tiene 47 puntos

$$\frac{37}{47} = 0.787$$

y la aproximación a π es 3.1489.

Primero observamos que conforme crece el número de puntos en una distribución uniforme dentro del cuadrado, la aproximación al valor de π se favorece. Con este método tan simple fácilmente obtuvimos tres dígitos del valor de π. Aunque el resultado rebasó su

valor. Esta es la desventaja con respecto al cálculo de π usando triángulos en un polígono dentro de un círculo de radio 1. La desventaja del método de triángulos es construir un polígono enorme que se aproxime al valor de π. Aunque, por otro lado, con el método de Montecarlo se necesita una base de datos muy grande para tener una mejor aproximación.

El método de Montecarlo (Sóbol, 1976) dio extrañamente un resultado que se aproximó rápidamente a π, y ahora se trata de aprovechar este camino con mayor tino o más eficientemente. Para esto lo primero que debemos observar es que la distribución uniforme de puntos se tomó aleatoriamente, esto es cada valor corresponde a un número aleatorio, o una variable aleatoria de la cual no conocemos el valor que tomará en un caso concreto dado, pero sabemos que valores podría tomar, y también las probabilidades de unos u otros valores. Con base en estos datos no se puede predecir con exactitud el resultado de una prueba relacionada con esta variable aleatoria, pero sí se puede prever con gran seguridad los resultados de un gran número de pruebas. En consecuencia, a estas variables aleatorias se les puede asociar directamente con la teoría de probabilidades. Una variable aleatoria ξ es discreta si puede tomar un conjunto discreto de valores $x_1, x_2, \dots, x_n$. Una variable aleatoria discreta ξ se define mediante la Tabla de *distribución de la variable aleatoria*

$$\xi = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$$

donde $x_1, x_2, \dots, x_n$ son los valores posibles de la variable ξ y $p_1, p_2, \dots, p_n$ son las probabilidades que les corresponden. La probabilidad de que la variable aleatoria ξ tome el valor x_i es igual a p_i:P$\{\xi = x_i\} = p_i$.

Los números $x_1, x_2, \dots, x_n$ pueden ser cualesquiera. En cambio, las probabilidades $p_1, p_2, \dots, p_n$ deben cumplir dos condiciones:

i) todos los números p_i deben ser positivos:

$$p_i > 0$$

ii) La suma de todos los p_i debe ser igual a 1.

$$p_1 + p_2 + \cdots + p_n = 1$$

Esta última condición significa que ξ necesariamente debe tomar en cada caso uno de los valores $x_1, x_2, \dots, x_n$.

Una variable aleatoria ξ es continua si puede tomar cualquier valor comprendido en un intervalo (a, b). Toda variable aleatoria continua ξ, queda definida si se da el intervalo

(a, b) que contiene los valores posibles de esta variable y la función $p(x)$ que lleva el nombre de *densidad de probabilidad* de la variable aleatoria ξ (o *densidad de distribución* de ξ).

La probabilidad de que ξ tome un valor perteneciente a un intervalo (a', b') [donde (a', b') es un intervalo cualquiera contenido en (a, b), es decir $a \leq a'$ y $b' \leq b$] es igual a la integral

$$\mathrm{P}\{a' < \xi < b'\} = \int_a^b p(x)dx.$$

El conjunto de los valores de ξ puede formar un intervalo cualquiera. En cambio, la densidad de probabilidad $p(x)$ debe cumplir las condiciones

i) la densidad $p(x)$ debe ser positiva

$$p(x) > 0$$

ii) la integral de la densidad $p(x)$ correspondiente a todo el intervalo (a, b) debe ser igual a 1:

$$\int_a^b p(x)dx = 1$$

En el esquema general de aplicación del método de Montecarlo supongamos que debemos calcular una magnitud *m* que se desconoce. Imaginemos una variable aleatoria ξ tal que su *esperanza* sea $\mathbf{M}\xi = m$ y su *varianza* sea $\mathbf{D}\xi = b^2$. Consideremos *N* variables aleatorias independientes $\xi_1, \xi_2, \ldots, \xi_n$ con la misma distribución que tiene ξ. Si *N* es suficientemente grande, la distribución de la suma $\rho_N = \xi_1 + \xi_2 + \cdots + \xi_n$ de acuerdo con el teorema del límite central, será una distribución aproximadamente normal con los parámetros $a = Nm$ y $\sigma^2 = Nb^2$. Esto es su esperanza y su varianza están dadas por

$$\mathbf{M}\xi = a \quad \text{y} \quad \mathbf{D}\xi = \sigma^2$$

Ejemplo: se define variable aleatoria *normal* ζ (o *gaussiana*) a aquella que está definida en todo el eje $(-\infty, \infty)$ y que tiene la densidad

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

donde a y $\sigma > 0$ son parámetros numéricos. El parámetro a no influye en la forma de la curva $p(x)$: su variación solo conduce a un desplazamiento de la curva a lo largo del eje *x*. En cambio, al variar σ si se altera la forma de la curva. Por ejemplo, si $x = a$

$$máx\ p(x) = p(a) = \frac{1}{\sigma\sqrt{2\pi}}$$

Si disminuimos σ, aumentará el $máx\ \ p(x)$. Pero toda el área que encierra la curva $p(x)$ es igual a 1 según la condición ii anterior. De manera que la curva tendrá que estirarse hacia arriba en una vecindad de $x = a$ y decrecer para todos los valores suficientemente grandes de *x*. Se puede demostrar que (Sóbol, 1976):

$$\int_{a-3\sigma}^{a+3\sigma} p\,(x)dx = 0.997$$

$$\mathrm{P}\{a - 3\sigma < \xi < a + 3\sigma\} = 0.997$$

$$\mathrm{P}\left\{Nm - 3b\sqrt{N} < \xi < Nm - 3b\sqrt{N}\right\} \approx 0.997$$

$$\mathrm{P}\left\{m - \frac{3b}{\sqrt{N}} < \frac{\xi}{N} < m - \frac{3b}{\sqrt{N}}\right\} \approx 0.997$$

$$\mathrm{P}\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\xi_j - m\right| < \frac{3b}{\sqrt{N}}\right\} \approx 0.997$$

Esta última relación ofrece un método para calcular *m* y estimar el error.

Integración Montecarlo

Considérese una función $g(x)$ definida en el intervalo $a < x < b$. En este punto, es preciso hallar el valor aproximado de la integral

$$I = \int_a^b g\,(x)dx$$

Tómese una densidad de distribución $p_\xi(x)$ cualquiera en el intervalo (a, b), que cumpla con las dos condiciones de la probabilidad para variables aleatorias continuas. Además de la variable aleatoria ξ definida con la densidad $p_\xi(x)$ en el intervalo (a, b), se necesita emplear la variable aleatoria

$$\eta = \frac{g(\xi)}{p_\xi(\xi)}$$

Esto es, la *esperanza* queda en la forma:

$$\mathbf{M}\eta = \int_a^b \left(\frac{g(x)}{p_\xi(x)}\right) p_\xi(x)dx = I$$

Considerando N variables aleatorias $\eta_1, \eta_2, \dots, \eta_N$ independientes e idénticas, y aplicando a la suma de estas variables el teorema del límite central, la última fórmula de la sección anterior ahora se puede escribir como

$$\mathrm{P}\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\eta_j - I\right| < 3\sqrt{\frac{\mathbf{D}\eta}{N}}\right\} \approx 0.997$$

Escogiendo N valores $\xi_1, \xi_2, \dots, \xi_N$, la última relación dice que para N suficientemente grande se tiene que

$$\frac{1}{N}\sum_{j=1}^{N}\eta_j \frac{g(\xi_j)}{p_\xi(\xi_j)} \approx I$$

Con gran probabilidad, el error de esta aproximación no pasa de

$$3\sqrt{\frac{\mathbf{D}\eta}{N}}$$

Además, una variable aleatoria η está *uniformemente distribuida en el intervalo* (a,b), si su densidad es constante en este intervalo:

$$p(x) = \frac{1}{b-a} \quad \text{para} \quad a < x < b$$

Para sortear los valores de ξ se tiene que

$$\int_a^{\xi} p\,(x)dx = \gamma$$

De tal manera que

$$\int_a^{\eta} \frac{1}{b-a}dx = \gamma$$

de donde

$$\eta = a + \gamma(b-a)$$

Entonces, para calcular la integral inicial se puede emplear una variable aleatoria ξ definida en el intervalo (a, b). En cualquier caso

$$\mathbf{M}\eta = \mathbf{M}\left[\frac{g(\xi)}{p_\xi(\xi)}\right] = I$$

Pero la varianza $\mathbf{D}\eta$ y la estimación del error dependen de que variable ξ se tome, ya que

$$\mathbf{D}\eta = \mathbf{M}(\eta^2) - I^2 = \int_a^b \left(\frac{g(x)}{p_\xi(x)}\right) p_\xi(x)dx - I^2$$

Se puede demostrar que esta expresión es mínima cuando $p_\xi(x)$ sea proporcional a $|g(x)|$. Por supuesto, no podemos escoger fórmulas muy complejas para $p_\xi(x)$ ya que resultaría laborioso sortear los valores de ξ.

Ejemplo numérico

Consideremos una función $f(x)>0$ definida en el intervalo (0,1)

$$f(x) = \sqrt{1-x^2}$$

Entonces, queremos aproximar el valor de la integral

$$I = 4\int_0^1 \sqrt{1-x^2}\,dx$$

El valor exacto de esta integral es

$$\begin{aligned}\int_0^1 \sqrt{1-x^2}\,dx &= \frac{1}{2}\left(x\sqrt{1-x^2} + arcsenx\right)_0^1 \\ &= \frac{1}{2}\left[\left(1\sqrt{1-1^2} + arcsen\ 1\right) - (0 + arcsen\ 0)\right] \\ &= \frac{1}{2}\left[\left(0 + \frac{\pi}{2}\right) - (0+0)\right] \\ &= \frac{\pi}{4} \\ \therefore\ I &= \pi\end{aligned}$$

Lo que sigue es obtener una aproximación al valor de π con el método de Montecarlo. Para esto se necesita al menos de una variable aleatoria ξ con una densidad constante $p_\xi(x) = \xi$, y otra con la densidad $g(x) = \sqrt{1-x^2}$. De tal manera que

$$I \approx \frac{1}{N}\sum_{j=1}^{N} \eta_j \frac{g(\xi_j)}{p_\xi(\xi_j)}$$

queda como

$$I \approx \frac{1}{N}\sum_{j=1}^{N} \eta_j \frac{\sqrt{1-\xi^2}}{\xi_j}$$

de la fórmula dada en la página previa, para el sorteo de ξ en el intervalo (0,1) se obtiene

$$\xi = 0 + \gamma(1-0) = \gamma$$

$$\therefore \quad \xi = \eta$$

con lo cual

$$I \approx \frac{1}{N}\sum_{j=1}^{N} \sqrt{1-\xi_j^2}$$

Además, se necesita generar números aleatorios en una distribución uniforme, útiles para este caso particular. Entre los números aleatorios más comunes están aquellos que se generan por sorteo, o con la ruleta, para los cuales deberá existir alguna fórmula adecuada para tomarlos. En este caso lo hemos hecho de manera directa arrojando disparos a gráficas como aquellas en la figura 50. Los datos generados aparecen en tres columnas en la tabla 1, una para cada inciso de la figura 50.

Tabla 1. *Parejas de valores obtenidos de disparo al blanco*

X	Y	X	Y	x	Y
0.24	0.68	0.32	0.50	0.42	0.77
0.40	0.26	0.06	0.81	0.28	0.81
0.52	0.52	0.48	0.39	0.10	0.69
0.14	0.23	0.04	0.04	0.65	0.56
0.83	0.67	0.62	0.02	0.39	0.60
0.12	0.39	0.65	0.18	0.91	0.59
0.49	0.13	0.72	0.81	0.73	0.28
0.84	0.09	0.84	0.25	0.42	0.03
0.78	0.15	0.91	0.45	0.29	0.38
0.31	0.12	0.49	0.65	0.18	0.12
0.6	0.7	0.75	0.56	0.50	0.80
0.71	0.41	0.61	0.28	0.01	0.35
0.34	0.97	0.20	0.01	0.11	0.51
0.32	0.25	0.15	0.94		
0.49	0.94	0.82	0.32		
0.63	0.97	0.88	0.05		
0.64	0.88	0.91	0.75		

Nota. La primera columna hacia la gráfica de la Figura 50a. La segunda columna hacia la gráfica de la Figura 50b. La tercera columna hacia la gráfica de la Figura 50c.

Fuente: elaboración propia.

Los valores de la coordenada x son los que se consideran para la variable aleatoria en este caso numérico. Tomando los primeros 10 valores de la columna de en medio se tiene que:

j	1	2	3	4	5	6	7	8	9	10
ξ_j	0.32	0.06	0.48	0.04	0.62	0.65	0.72	0.84	0.91	0.49
$\sqrt{1-\xi_j^2}$	0.95	0.998	0.88	0.999	0.78	0.76	0.69	0.54	0.41	0.87

De donde

$$I \approx \frac{1}{10}(7.88)$$

Multiplicando por 4 para completar los cuatro cuadrantes del plano cartesiano $I \approx 4(0.788)$ se obtiene la siguiente aproximación a π:

$$I \approx 3.15149$$

Si escogemos aleatoriamente 10 datos de la Tabla 1, se tiene que

j	1	2	3	4	5	6	7	8	9	10
ξ_j	0.52	0.75	0.10	0.72	0.48	0.64	0.34	0.49	0.73	0.84
$\sqrt{1-\xi_j^2}$	0.85	0.66	0.99	0.69	0.88	0.77	0.94	0.87	0.68	0.54

de donde $I \approx 4(0.787)$,

$$I \approx 3.148$$

resultó ser una elección afortunada que produjo tres dígitos de π en error de truncamiento. Usando los 47 datos en un cálculo similar, el valor aproximado a π es

$$I \approx 3.206$$

Consideremos ahora la evaluación de la integral configuracional de un ensamble canónico en mecánica estadística.

$$Z_{NVT} = \int d\boldsymbol{r}\, exp(-\beta V)$$

para un sistema de 100 moléculas (N=100) en un cubo de lado L. Una integración por regla de Simpson podría requerir 10 evaluaciones de la función para cada una de las 300 coordenadas y, además, recorrer el intervalo $\left(-\frac{1}{2}L, \frac{1}{2}L\right)$. Un total de 10^{300} evaluaciones de

la función, lo cual es enorme e inviable. Sin embargo, usando el método de Montecarlo se puede calcular la energía potencial de cada configuración del sistema, junto con los valores de otras propiedades directamente de las posiciones de los átomos. El método de Montecarlo muestrea las posiciones de las partículas de un espacio 3N-dimensional. Reescribiendo la función de partición del ensamble canónico para un sistema de *N* partículas idénticas de masa *m*,

$$\boldsymbol{Q}_{NVT} = \frac{1}{N!}\frac{1}{h^{3N}}\iint d\boldsymbol{p}^N d\boldsymbol{r}^N exp\left[-\frac{H(\boldsymbol{p}^N, \boldsymbol{r}^N)}{k_B T}\right]$$

Donde el hamiltoniano que corresponde a la energía total del sistema está dado por

$$H(\boldsymbol{p}^N, \boldsymbol{r}^N) = \sum_{i=1}^{N} \frac{|\boldsymbol{p}_i|^2}{2m} + V(\boldsymbol{r}^N),$$

el factor *N*! desaparece cuando las partículas ya no son indistinguibles. El valor del hamiltoniano depende de las 3*N* posiciones y 3*N* momentos de las partículas en el sistema. El punto crucial por reconocer es que la doble integral se puede separar en dos integrales separadas, una sobre las posiciones y la otra sobre los momentos:

$$\boldsymbol{Q}_{NVT} = \frac{1}{N!}\frac{1}{h^{3N}}\int d\boldsymbol{p}^N\, exp\left[-\frac{|\boldsymbol{p}|^2}{2mk_B T}\right]\int d\boldsymbol{r}^N\, exp\left[-\frac{V(\boldsymbol{r}^N)}{k_B T}\right]$$

Esta separación es posible si la función de energía potencial $V(\boldsymbol{r}^N)$ no depende de las velocidades. De esta manera se puede efectuar analíticamente la integral sobre los momentos, esto es:

$$\int d\boldsymbol{p}^N\, exp\left[-\frac{|\boldsymbol{p}|^2}{2mk_B T}\right] = (2\pi mk_B T)^{3N/2}$$

Entonces la función de partición se puede reescribir como

$$\boldsymbol{Q}_{NVT} = \frac{1}{N!}\left(\frac{2\pi mk_B T}{h^2}\right)^{3N/2}\int d\boldsymbol{r}^N\, exp\left[-\frac{V(\boldsymbol{r}^N)}{k_B T}\right]$$

y la integral sobre las posiciones es la integral configuracional Z_{NVT}:

$$Z_{NVT} = \int d\boldsymbol{r}^N\, exp\big(-\beta V(\boldsymbol{r}^N)\big)$$

En un gas ideal no hay interacciones entre las partículas, entonces $V(\boldsymbol{r}^N) = 0$ y $exp\big(-\beta V(\boldsymbol{r}^N)\big)$ es igual a 1 para cada partícula de gas en el sistema. La integral de 1 sobre las coordenadas de cada átomo es igual al volumen, y así para *N* partículas de gas ideal la

integral configuracional está dada por V^N. Esto asigna el siguiente resultado para la función de partición canónica de un gas ideal:

$$\boldsymbol{Q}_{NVT} = \frac{V^N}{N!}\left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2}$$

y es reescrita en términos de la longitud de onda *térmica de DeBroglie* Λ:

$$\boldsymbol{Q}_{NVT} = \frac{V^N}{N!\,\Lambda^{3N}}$$

$$\text{con } \Lambda = \sqrt{h^2/2\pi m k_B T}.$$

Se puede ver entonces que la función de partición para un sistema real tiene una contribución debida a la conducta de gas ideal y una contribución debida a las interacciones entre las partículas. Entonces la función de partición queda:

$$\boldsymbol{Q}_{NVT} = \boldsymbol{Q}_{NVT}^{ideal}\boldsymbol{Q}_{NVT}^{exceso}$$

La parte de exceso de la función de partición está dada por

$$\boldsymbol{Q}_{NVT}^{exceso} = \frac{1}{V^N}\int d\boldsymbol{r}^N \, exp\left[-\frac{V(\boldsymbol{r}^N)}{k_B T}\right]$$

Las contribuciones de gas ideal se pueden determinar analíticamente integrando sobre los momentos. Por ejemplo, la energía libre de Helmholtz está relacionada a la función de partición canónica por:

$$\begin{aligned} A &= -k_B T ln \boldsymbol{Q}_{NVT} \\ &= -k_B T ln \boldsymbol{Q}_{NVT}^{ideal}\boldsymbol{Q}_{NVT}^{exceso} \end{aligned}$$

Por lo tanto

$$A = A^{ideal} + A^{exceso}$$

Lo relevante es que todas las desviaciones de la conducta de gas ideal se deben a la presencia de interacciones entre los átomos en el sistema, conforme se calculan usando la función de energía potencial. Esta función de energía depende solo de las posiciones de los átomos y no de sus momentos, y así una simulación Montecarlo puede calcular las contribuciones de exceso que producen desviaciones de la conducta de gas ideal.

Por ejemplo, la energía potencial promedio se puede determinar evaluando la integral

$$\langle V(\boldsymbol{r}^N)\rangle = \int d\boldsymbol{r}^N \, V(\boldsymbol{r}^N)\rho(\boldsymbol{r}^N)$$

con

$$\rho(\boldsymbol{r}^N) = \frac{exp[-\beta V(\boldsymbol{r}^N)]}{Z}$$

El denominador Z es la integral configuracional. Para una función potencial comúnmente usada en modelamiento molecular no es posible evaluar estas integrales analíticamente. Sin embargo, se puede intentar obtener valores para las integrales usando métodos numéricos, lo cual lleva a cálculos inviables. Para calcular la función de partición para un sistema de *N* átomos usando el método de integración de Montecarlo se pueden usar los siguientes pasos:

1. Obtener una configuración del sistema generando aleatoriamente 3*N* coordenadas Cartesianas, que son asignadas a las partículas.
2. Calcular la energía potencial de la configuración $V(\boldsymbol{r}^N)$.
3. De la energía potencial, calcular el factor de Boltzmann, $exp[-V(\boldsymbol{r}^N)/k_BT]$
4. Sumar el factor de Boltzmann a la suma acumulada de factores de Boltzmann y la contribución de la energía potencial a su suma acumulada, y regresar al paso 1.
5. Después de un número N_{prueba} de interacciones, el valor medio de la energía potencial se calcula usando:

$$\langle V(\boldsymbol{r}^N)\rangle = \frac{\sum_{i=1}^{N_{prueba}} V_i(\boldsymbol{r}^N) exp[-V_i(\boldsymbol{r}^N)/k_BT]}{\sum_{i=1}^{N_{prueba}} exp[-V_i(\boldsymbol{r}^N)/k_BT]}$$

O en general para cualquier propiedad

$$\langle A\rangle_{NVT} = \frac{\int d\boldsymbol{r}\, A exp(-\beta V)}{\int d\boldsymbol{r}\, exp(-\beta V)} \approx \frac{\sum_{\tau=1}^{\tau_{max}} A(\tau) exp[-\beta V(\tau)]}{\sum_{\tau=1}^{\tau_{max}} exp[-\beta V(\tau)]}$$

Desafortunadamente esta no es una aproximación viable para calcular propiedades termodinámicas debido al gran número de configuraciones que tienen factores de Boltzmann extremadamente pequeños causados por traslapes de alta-energía entre las partículas. Esto refleja la naturaleza del espacio fase, que corresponde a configuraciones no-físicas con muy altas energías. Las regiones de baja energía coinciden con fases observadas físicamente, tales como sólido, líquido, etc.

Importancia del muestreo

Existen muchas maneras de generar números aleatorios, pero solo con la práctica se puede dilucidar si son útiles para describir el objetivo buscado. Además, dicha descripción siempre es aproximada. Algunos ejemplos de obtención de variables aleatorias son los siguientes:

i) Escribamos los números 0, 1, 2, … , 9, cada uno en una tarjeta, considerando que las 10 tarjetas son iguales. Metamos todas las tarjetas en un sombrero y, después de mezclarlas bien, comencemos a extraerlas, devolviendo al sombrero la tarjeta sacada para mezclar de nuevo todas las tarjetas. Las cifras así obtenidas se pueden recopilar en una tabla. Esta tabla se conoce como *tabla de números aleatorios* y se puede introducir en la memoria de una computadora. Cada variable aleatoria tiene una probabilidad 0.1, y en total suman 1.

ii) Una tabla muy grande de números aleatorios se construye con una ruleta electrónica. Es preciso subrayar que, aun así, no es tan fácil obtener una buena tabla de números aleatorios. Todo aparato físico real genera variables aleatorias cuyas distribuciones difieren ligeramente de la distribución ideal. Además, durante el experimento pueden surgir errores. Por eso, las tablas obtenidas suelen comprobarse minuciosamente a base de pruebas estadísticas especiales, para ver si no hay contradicción entre unas u otras propiedades de grupos de números, y la hipótesis de que estos números representan valores de la variable aleatoria.

Veamos una de las *pruebas* más sencillas. Consideremos una tabla que contiene N cifras. Supongamos que el número de ceros en esta tabla es ν_0, el número de unos es ν_1, el número de doses es ν_2, etc. Calculemos la suma

$$\sum_{i=0}^{9}(\nu_i - 0.1N)^2$$

La teoría de probabilidades permite predecir los límites entre los cuales puede estar comprendida esta suma que no puede ser excesivamente grande (por cuanto la esperanza matemática de cada ν_i es igual a $0.1N$) ni excesivamente pequeña (ya que esto equivaldría a una *regularidad excesiva* en la distribución de los valores).

Las tablas de números aleatorios se emplean más comúnmente cuando los cálculos correspondientes al método de Montecarlo se realizan a mano. Si la capacidad de memoria en una computadora es considerablemente grande, entonces las tablas de números aleatorios resultan de gran utilidad.

iii) Todo dispositivo mecánico resulta demasiado lento comparado con una computadora. Por eso para generar variables aleatorias suelen emplearse los ruidos en las lámparas electrónicas: si durante un intervalo fijo Δt de tiempo el nivel de ruido sobrepasa el

umbral escogido un número par de veces, se inscribe el cero; si esto ocurre un número impar de veces, se inscribe el uno.

Imaginemos *m* generadores que funcionan paralelamente sin cesar, y envían ceros y unos aleatorios a todos los órdenes binarios de una célula especial. Cada ciclo corresponderá a un número de *m* órdenes. Durante el cálculo podemos dirigirnos en cualquier momento a esta célula y tomar de ella un valor de la variable aleatoria γ uniformemente distribuida en el intervalo (0,1). Este valor será aproximado y quedará representado por una fracción binaria de m órdenes $0, \alpha_{(1)}\alpha_{(2)} \ldots \alpha_{(m)}$, donde cada $\alpha_{(i)}$ imita la variable aleatoria con la distribución:

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Sin embargo, también este método tiene sus defectos. En primer lugar, es difícil comprobar la *calidad* de los números aleatorios generados. Se hace necesario un control periódico ya que debido a distintos desarreglos puede surgir la así llamada *deriva en la distribución* (cuando los ceros y los unos en uno de los órdenes dejan de aparecer con la misma frecuencia). En segundo lugar, todos los cálculos en las computadoras suelen realizarse dos veces para evitar intermitencias casuales. Pero, sin recurrir a la memoria durante los cálculos, es imposible reproducir los mismos números aleatorios, y si recurrimos a la memoria regresamos al caso de las tablas.

iv) Técnica de aceptación o rechazo. Existe un gran número de métodos adecuados para generar un vector en la superficie de una esfera unitaria. El más simple usa la técnica de aceptación o rechazo. El procedimiento es iterativo:

a) Generar tres variables aleatorias uniformes, $\xi_1, \xi_2, \text{y } \xi_3, \text{en } (0,1)$.

b) Calcular $\zeta_i = 1 - 2\xi_1$, para *i*=1,3 tal que el vector $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \zeta_3)$ se distribuya en un cubo de lado 2, centrado en el origen.

c) Formar la suma $\zeta^2 = \zeta_1^2 + \zeta_2^2 + \zeta_3^2$.

d) Para $\zeta^2 < 1$ (*i.e.*, inscrita en el interior de la esfera) tomar $\hat{\boldsymbol{\zeta}} = (\zeta_1/\zeta\,, \zeta_2/\zeta\,, \zeta_3/\zeta)$ como el vector.

e) Para $\zeta^2 > 1$, rechazar el vector y regresar al paso (a).

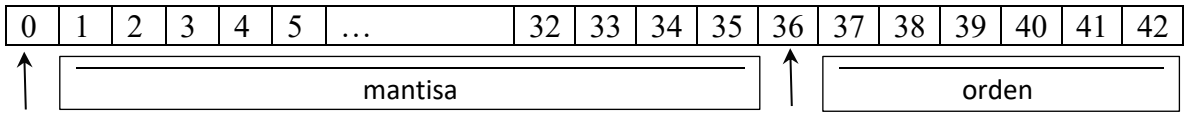
La eficiencia de este método tiende a $\pi/6$ tal que el algoritmo requiere 5.73 variables uniformes en promedio.

v) Los números que se obtienen a partir de una fórmula y que imitan los valores de la variable aleatoria γ se denominan *números pseudoaleatorios*. El primer algoritmo para construirlos se conoce como *método de los centros cuadrados*. Como ejemplo consideremos el número $\gamma_0 = 0.9876$ formado por cuatro cifras. Al elevarlo al cuadrado se obtiene el número $\gamma^2 = 0.97535376$, de ocho cifras. Tomemos las cuatro cifras que aparecen en el centro de este número y pongamos $\gamma_1 = 0.5353$. Elevando ahora γ_1 al cuadrado $(\gamma_1^2 = 0.28654609)$, escojamos de nuevo las cuatro cifras del centro y pongamos $\gamma_2 = 0.6546$. Procediendo de la misma forma tendremos $\gamma_2^2 = 0.42850116$; $\gamma_3 = 0.8501$, $\gamma_3^2 = 0.72267001$; $\gamma_4 = 0.2670$, $\gamma_4^2 = 0.07128900$; $\gamma_5 = 0.1289$, etc.

Este método no puede considerarse satisfactorio pues da más números pequeños de lo que se necesita. Por eso se han desarrollado otros algoritmos. Uno de estos se conoce como Strelá.

Strelá es una computadora de tres direcciones con coma flotante. La célula en la que se inscribe el número x consta de 43 órdenes binarios (ver figura 51). La computadora opera con números binarios en forma normalizada: $x = \pm q \cdot 2^{\pm p}$, donde p es el orden y q es la mantisa del número. En el orden j-ésimo de la célula puede figurar un cero o un uno.

Figura 51. *Celda de 43 órdenes binarios*



Fuente: Sóbol (1976).

Representemos, por ejemplo, esta variable, entonces

$$q = \frac{\epsilon_1}{2^1} + \frac{\epsilon_2}{2^2} + \cdots + \frac{\epsilon_{35}}{2^{35}}$$

$$p = \epsilon_{37}2^5 + \epsilon_{38}2^4 + \cdots + \epsilon_{42}2^0$$

La condición de normalización $0 \leq q < 1$ significa que ϵ_1 siempre es igual a 1. El signo «+» se representa por el cero y el signo «-» por el uno.

El número γ_{k+1} se obtiene del número γ_k realizando tres operaciones:

1) γ_k se multiplica por una constante grande que, como regla, se toma igual a 10^{17};
2) la imagen del producto $10^{17}\gamma_k$ se desplaza en siete órdenes hacia la izquierda (de modo que los siete primeros órdenes del producto desaparecen, mientras que en los órdenes del 36 al 42 aparecen ceros); y
3) se toma el valor absoluto del número obtenido (después de normarlo) que será precisamente γ_{k+1}.

Si arrancamos de $\gamma_0 = 1$, este proceso permite obtener más de 80 000 números γ_k diferentes; después en la sucesión surge un periodo y los números comienzan a repetirse. Diferentes controles de los primeros 50 000 números han dado resultados plenamente satisfactorios. Estos números se han empleado más de una vez para la solución de diversos problemas.

Aun se puede hablar del sorteo de una variable discreta, del sorteo de una variable continua, etc. Sin embargo, ahora vamos a introducir a las cadenas de Markov ya que se requieren para encontrar propiedades termofísicas. Por ejemplo: sea $q(\boldsymbol{x} \rightarrow \boldsymbol{x}')$ la probabilidad de que el proceso haga una transición del estado $\boldsymbol{x}$ al estado $\boldsymbol{x}'$

Esta probabilidad de transición define lo que es conocido como *cadena de Markov* en el espacio de estados. La cadena de Markov definida por q debe ser ergódica, esto es que esencialmente cada estado debe ser alcanzable desde cualquier otro y puede no haber estrictamente ciclos periódicos.

Supongamos que corremos la cadena de Markov para t pasos, y sea $\Pi_t(\boldsymbol{x})$ la probabilidad de que el sistems este en el estado $\boldsymbol{x}$ en el tiempo t. Análogamente, sea $\Pi_{t+1}(\boldsymbol{x}')$ la probabilidad de estar en el estado $\boldsymbol{x}'$ en el tiempo t+1. Dada $\Pi_t(\boldsymbol{x})$, podemos calcular $\Pi_{t+1}(\boldsymbol{x}')$ sumando para todos los estados la probabilidad de estar en ese estado veces la probabilidad de hacer la transición a $\boldsymbol{x}'$:

$$\Pi_{t+1}(\boldsymbol{x}') = \sum_{\boldsymbol{x}} \Pi_t(\boldsymbol{x})\, q(\boldsymbol{x} \rightarrow \boldsymbol{x}')$$

La cadena alcanza su distribución estacionaria si $\Pi_t = \Pi_{t+1}$. Llamemos Π a esta distribución estacionaria, entonces la ecuación que la define es

$$\Pi(\boldsymbol{x}') = \sum_{\boldsymbol{x}} \Pi(\boldsymbol{x})\, q(\boldsymbol{x} \rightarrow \boldsymbol{x}') \qquad \forall \boldsymbol{x}'$$

Bajo ciertas suposiciones estándar acerca de la distribución de probabilidades de transición q, existe exactamente una distribución Π que satisface esta ecuación para cualquier q dada.

Esta ecuación se puede leer como diciendo que el *flujo de salida* esperado de cada estado (*i.e.* su *población* común) es igual al *flujo de entrada* esperado de todos los estados. Una manera obvia de satisfacer esta relación es si el flujo esperado entre cualquier par de estados es el mismo en ambas direcciones. Esta es la propiedad de *balance detallado*:

$$\Pi(\boldsymbol{x})\, q(\boldsymbol{x} \to \boldsymbol{x}') = \Pi(\boldsymbol{x}')\, q(\boldsymbol{x}' \to \boldsymbol{x}) \qquad \forall \boldsymbol{x}, \boldsymbol{x}'$$

Se puede demostrar que *balance detallado* implica estacionariedad, simplemente sumando sobre $\boldsymbol{x}$:

$$\sum_{\boldsymbol{x}} \Pi(\boldsymbol{x})\, q(\boldsymbol{x} \to \boldsymbol{x}') = \sum_{\boldsymbol{x}} \Pi(\boldsymbol{x}')\, q(\boldsymbol{x}' \to \boldsymbol{x}) = \Pi(\boldsymbol{x}') \sum_{\boldsymbol{x}} q\,(\boldsymbol{x}' \to \boldsymbol{x}) = \Pi(\boldsymbol{x}')$$

donde el último paso se sigue debido a que se garantiza que ocurre una transición de $\boldsymbol{x}'$.

Las técnicas de muestreo para elegir números aleatorios de una distribución $\rho(x)$ permiten que la evaluación de la función esté concentrada en las regiones del espacio que hacen contribuciones importantes a la integral.

Consideremos el ensamble canónico. En este caso la integral deseada es

$$\langle A \rangle_{NVT} = \int d\Gamma\, \rho_{NVT}(\Gamma) A(\Gamma)$$

esto es, el integrando es $\rho = \rho_{NVT} A$. Muestreando configuraciones al azar de una distribución ρ elegida se puede estimar la integral como

$$\langle A \rangle_{NVT} = \langle A\, \rho_{NVT} / \rho \rangle_{pruebas}$$

Esto no siempre es verdadero y algunas veces elegimos distribuciones alternativas de $\rho(\Gamma)$. Dicho método, con $\rho = \rho_{NVT}$, originalmente fue desarrollado por Metropolis et al. (1953). El problema no está resuelto, simplemente está refraseado. La dificultad del trabajo es encontrar un método para generar una secuencia de estados al azar tal que al final de la simulación cada estado haya ocurrido con la probabilidad adecuada. Resulta que es posible hacer esto sin calcular el factor de normalización para ρ_{NVT}, *i.e.* la función de partición.

La solución es establecer una cadena de Markov de estados del líquido, construida de tal manera que tenga una distribución limitante de ρ_{NVT}. Una cadena de Markov es una secuencia de pruebas que satisface dos condiciones:

a) El resultado de cada prueba pertenece a un conjunto finito de resultados $\{\Gamma_1, \Gamma_2, \dots, \Gamma_m, \Gamma_n, \dots\}$, llamado el espacio de estado.

b) El resultado de cada prueba depende solo del resultado de la prueba que le precede inmediatamente.

Dos estados Γ_m y Γ_n están ligados por una probabilidad de transición Π_{mn} , que es la probabilidad de ir de un estado *m* a un estado *n*. Las propiedades de una cadena de Markov se ilustran mejor con un simple ejemplo. Suponga que la confiabilidad de su computadora sigue cierto patrón. Si está encendida y corriendo en un día tiene un 60 por ciento de oportunidad de correr correctamente en el siguiente. Sin embargo, si está apagada, tiene un 70 por ciento de oportunidad de estar sin correr el día siguiente. El espacio de estado tiene dos componentes, arriba (↑) y abajo (↓), y la matriz de transición tiene la forma

$$\pi = \begin{matrix} \\ \uparrow \\ \downarrow \end{matrix} \begin{matrix} \begin{matrix} \uparrow & \downarrow \end{matrix} \\ \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix} \end{matrix}$$

Si es igualmente probable que la computadora esté encendida o apagada al comenzar, entonces la probabilidad inicial se puede representar como un vector, que tiene las dimensiones del espacio de estado

$$\begin{matrix} \uparrow & \downarrow \end{matrix}$$
$$\rho^{(1)} = (0.5 \ , \ 0.5)$$

La probabilidad de que la computadora este encendida en el segundo día está dada por la ecuación matricial

$$\rho^{(2)} = \rho^{(1)}\pi = (0.45 \ , \ 0.55)$$

Es decir, existe un 45 % de oportunidad de correr un programa. El siguiente día daría:

$$\rho^{(3)} = \rho^{(2)}\pi = \rho^{(1)}\pi\pi = \rho^{(1)}\pi^2 = (0.435 \ , \ 0.565)$$

Y hay 43.5 % de oportunidad de éxito. Para calcular las oportunidades en una corrida larga, la distribución limitante está dada por

$$\rho = \lim_{\tau \to \infty} \rho^{(1)}\pi^{\tau}$$

Unas cuantas aplicaciones de esta ecuación muestran que el resultado converge a $\rho = (0.4286, 0.5714)$. Está claro de esta ecuación que la distribución limitante, ρ, debe satisfacer la ecuación de eigenvalores

$$\left.\begin{aligned} \rho\pi &= \rho \\ \sum_m \rho_m \pi_{mn} &= \rho_n \end{aligned}\right\} \tag{93}$$

Con eigenvalor unitario. A π se le denomina matriz estocástica, ya que sus filas suman 1.

$$\sum_m \pi_{mn} = 1 \tag{94}$$

Es la matriz de transición para una cadena de Markov irreducible (una cadena ergódica o irreducible es aquella en la que eventualmente cada estado se puede alcanzar desde otro estado). Mas formalmente, notamos que el teorema de Perron-Frobenius dice que una matriz estocástica irreducible ha dejado un eigenvalor que es igual a la unidad, y el eigenvector correspondiente es la distribución limitante de la cadena. Los otros eigenvalores son menores a la unidad y gobiernan la tasa de convergencia de la cadena de Markov. La distribución limitante ρ implicada por la cadena es completamente independiente de la condición inicial $\rho^{(1)}$ (no hay que preocuparse de que hoy esté probablemente apagada).

En el caso de un líquido debemos construir una matriz de transición mucho más grande, que sea estocástica y ergódica. En contraste al problema previo, los elementos de la matriz de transición son desconocidos, pero la distribución limitante de la cadena es el vector con elementos $\rho_m = \rho_{NVT}(\Gamma_m)$ para cada punto en el espacio fase. Es posible determinar elementos de π que satisfagan las ecuaciones (93) y (94), y así generar una trayectoria de espacio fase en el ensamble canónico. Tenemos libertad considerable en encontrar una matriz de transición apropiada, con la restricción crucial de que los elementos de la matriz deberían ser independientes de Q_{NVT}. Un truco útil en buscar una solución de la ecuación (94) es reemplazarla por la condición innecesariamente fuerte de 'reversibilidad microscópica':

$$\rho_m \pi_{mn} = \rho_n \pi_{nm} \tag{95}$$

Al sumar sobre todos los estados *m* y hacer uso de la ecuación (94) recuperamos la ecuación (93)

$$\sum_m \rho_m \pi_{mn} = \sum_m \rho_n \pi_{nm} = \rho_n \sum_m \pi_{nm} = \rho_n$$

Un esquema adecuado para construir una trayectoria de espacio fase en el ensamble canónico involucra elegir una matriz de transición que satisfaga las ecuaciones (94) y (95). El primero de tales esquemas fue sugerido por Metropolis et al. (1953) y frecuentemente es conocido como la solución asimétrica. Si los estados *m* y *n* son distintos, esta solución considera dos casos

$$\pi_{mn} = \alpha_{mn} \qquad \rho_n \geq \rho_m \qquad m \neq n \tag{96}$$

$$\pi_{mn} = \alpha_{mn}\,(\rho_n/\rho_m) \qquad \rho_n < \rho_m \qquad m \neq n \tag{97}$$

También es importante permitir la probabilidad de que el líquido permanezca en el mismo estado:

$$\pi_{mm} = 1 - \sum_{n \neq m} \pi_{mn} \tag{98}$$

En esta solución α es una matriz estocástica simétrica ($\alpha_{mn} = \alpha_{nm}$) a menudo llamada la matriz subyacente de la cadena de Markov. Las propiedades simétricas de α pueden ser usadas para mostrar que en los tres casos ($\rho_m = \rho_n$, $\rho_m < \rho_n$ y $\rho_m > \rho_n$) la matriz de transición definida en (96-98) satisface las ecuaciones (94) y (95). Vale la pena destacar que es la propiedad simétrica de α la que es esencial en satisfacer reversibilidad microscópica en este caso. Se pueden construir matrices α no-simétricas que satisfagan reversibilidad microscópica o solo la condición más débil, ecuación (93), pero estas no forman parte de la receta básica de Metropolis. Finalmente, esta solución solo involucra la tasa ρ_n/ρ_m y por lo tanto es independiente de Q_{NVT}.

Existen otras soluciones a las ecuaciones (94) y (95). La solución simétrica a menudo es referida como muestreo de Barker.

$$\pi_{mn} = \alpha_{mn}\rho_n/(\rho_n + \rho_m) \qquad m \neq n \tag{99}$$

$$\pi_{mm} = 1 - \sum_{n \neq m} \pi_{mn} \tag{100}$$

La condición (99-100) también satisface la reversibilidad microscópica.

Si se generan estados del fluido usando matrices de transición tales como las ecuaciones (99-100), entonces una propiedad particular, $\langle A \rangle_{run}$, obtenida al promediar sobre

las pruebas τ_{run} en la cadena de Markov está relacionada al promedio en el ensamble canónico.

$$\langle A\rangle_{NVT} = \langle A\rangle_{run} + \mathcal{O}\left(\tau_{run}^{-\frac{1}{2}}\right) \tag{101}$$

Puesto que hay un número de matrices de transición adecuado, es útil elegir una solución particular que minimice la varianza en la estimación de $\langle A\rangle_{run}$. En particular la *ineficiencia estadística*

$$s = \lim_{\tau_{run}\to\infty} \tau_{run}\, \sigma^2\left(\langle A\rangle_{run}\right)/\sigma^2(A) \tag{102}$$

mide qué tan lentamente converge una corrida a su valor limitante. Se ha demostrado que para ordenar dos matrices de transición es razonable

$$\pi_1 \leq \pi_2$$

Si cada elemento fuera de la diagonal de π_1 sería menor que el elemento correspondiente en π_2, entonces

$$s(\langle A\rangle, \pi_1) \geq s(\langle A\rangle, \pi_2)$$

Para cualquier propiedad A, si los elementos fuera de la diagonal de π son grandes, entonces la probabilidad de quedar en el mismo estado es pequeña, y el muestreo del espacio fase será mejorado. Con la restricción de que ρ_m y ρ_n sean positivas, las ecuaciones (96-100) muestran que la solución de Metropolis conduce a una ineficiencia estadística más baja de la media que la solución de Barker. Una ineficiencia estadística baja no es el único criterio para elegir una π particular. Dado que las simulaciones son de longitud finita, es esencial que la cadena de Markov muestree una porción representativa del espacio fase en un número razonable de movimientos. Todos los resultados derivados en esta sección dependen de la ergodicidad de la cadena (*i.e.* que hay alguna probabilidad de transición multi-paso distinta de cero, de moverse entre dos estados cualesquiera permitidos del fluido). Si estos estados permitidos no están conectados, la corrida MC podría producir un *s* bajo, pero en adición una estimación pobre del promedio canónico. Cuando la trayectoria entre dos regiones permitidas del espacio fase es difícil de encontrar, la situación se describe como un cuello de botella. Estos cuellos de botella siempre han sido preocupantes en simulaciones MC, pero son particularmente problemáticos en la simulación de coexistencia de dos fases, y en la

simulación de líquidos ordinarios en densidad usualmente alta. Donde se ha hecho una comparación entre las dos soluciones comunes a la matriz de transición, ecuaciones (96-100), la solución de Metropolis parece llevar a una convergencia más rápida de la cadena.

El método de Metropolis llega a ser más favorable conforme el número de estados disponibles incrementa en un paso dado, y conforme incrementa la diferencia de energía entre los estados.

El método de Metropolis

En una simulación Montecarlo las configuraciones sucesivas del sistema no están conectadas en el tiempo. Más bien, cada configuración depende solo de su predecesor y no de ninguna otra de las configuraciones visitadas previamente. El método Montecarlo genera configuraciones al azar y utiliza un conjunto especial de criterios para decidir si acepta o no cada nueva configuración. Estos criterios aseguran que la probabilidad de obtener una configuración dada es igual a su factor de Boltzmann $exp\{-V(\boldsymbol{r}^N)/k_BT\}$, donde $V(\boldsymbol{r}^N)$ se calcula utilizando la función de energía potencial. De modo que estados con una energía baja, se generan con mayor probabilidad que las configuraciones con una energía más alta. Para cada configuración que se acepta, se calculan los valores de las propiedades deseadas y al final del cálculo se obtiene el promedio de estas propiedades simplemente promediando el número de valores calculados, M:

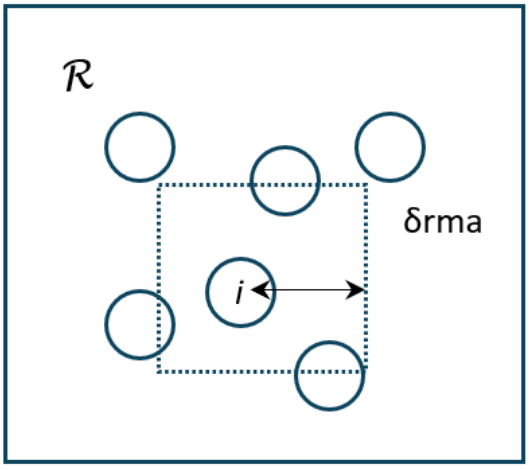
$$\langle A\rangle = \frac{1}{M}\sum_{i=1}^{M} A(\boldsymbol{r}^N)$$

Metropolis et al. (1953) reportaron el primer cálculo tipo Montecarlo. No obstante, existen otras formas en que se puede generar un conjunto de configuraciones. Entonces, el esquema de Metropolis es solo una de varias posibilidades, aunque es, con mucho, el más popular. En una simulación de Montecarlo, cada nueva configuración del sistema puede generarse moviendo aleatoriamente un solo átomo o molécula. En algunos casos, también se pueden obtener nuevas configuraciones moviendo varios átomos o moléculas o girando alrededor de uno o más enlaces. La energía de la nueva configuración se calcula utilizando la función de energía potencial. Si la energía de la nueva configuración es menor que la energía de su predecesor, entonces se acepta la nueva configuración. Si la energía de la nueva configuración es mayor que la energía de su predecesor, entonces el factor de Boltzmann de

la diferencia de energía se calcula $exp\left[-\left(V_{new}(\boldsymbol{r}^N) - V_{old}(\boldsymbol{r}^N)\right)/k_B T\right]$. Luego se genera un número aleatorio entre 0 y 1, y se compara con este factor de Boltzmann. Si el número aleatorio es mayor que el factor de Boltzmann, el movimiento se rechaza y la configuración original se conserva para la próxima iteración; si el número aleatorio es menor, se acepta el movimiento y la nueva configuración se convierte en el siguiente estado. Este procedimiento tiene el efecto de permitir movimientos a estados de mayor energía. Cuanto más pequeño sea el movimiento ascendente (es decir, cuanto menor sea el valor de $V_{new}(\boldsymbol{r}^N) - V_{old}(\boldsymbol{r}^N)$ mayor será la probabilidad de que el movimiento sea aceptado.

Para implementar la solución de Metropolis a la matriz de transición es necesario especificar la matriz α estocástica subyacente. La matriz está diseñada para llevar el sistema desde el estado *m* en cualquiera de sus estados circundantes *n* con igual probabilidad. Existe una libertad considerable para elegir α, y la única constricción es que $\alpha_{mn} = \alpha_{nm}$. Un ejemplo de un estado circundante se ilustra en la figura 52.

Figura 52. *Estado circundante*



Fuente: elaboración propia.

Se muestran seis átomos en un estado *m*; para construir un estado circundante *n*, se elige un átomo (*i*) al azar y se desplaza de su posición r_i^m con igual probabilidad a cualquier punto r_i^n interior del cuadro $\mathcal{R}$. Este cuadro es de lado $2\delta r_{max}$ y se centra en r_i^m. Tridimensionalmente serían esferas dentro de un cubo. En una computadora se pueden

manejar los pixeles para un número finito pero grande de posiciones $N_{\mathcal{R}}$ para el átomo *i* y en este caso α_{mn} se puede definir como

$$\alpha_{mn} = {}^{1}/_{N_{\mathcal{R}}} \qquad r_i^n \in \mathcal{R}$$

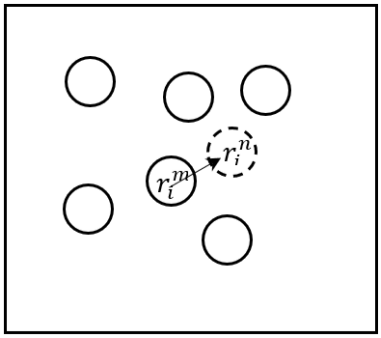
$$\alpha_{mn} = 0 \qquad r_i^n \notin \mathcal{R}$$

Con esta elección de α, la ecuación (96-98) se implementa fácilmente. Al comienzo de una corrida, MC mueve un átomo que se recoge aleatoriamente y se le da un desplazamiento aleatorio uniforme a lo largo de cada una de las direcciones de las coordenadas.

El máximo desplazamiento δr_{max} es un parámetro ajustable que gobierna el tamaño de la región $\mathcal{R}$ y controla la convergencia de la cadena de Markov. La nueva posición se obtiene con un código de *random* (drmax es el desplazamiento máximo δr_{max}, rxipost = rx(i) + (2.0*ranf(dummy)-1.0)*drmax, ryipost = ry(i) + (2.0*ranf(dummy)-1.0)*drmax, rzipost = rz(i) + (2.0*ranf(dummy)-1.0)*drmax).

El elemento apropiado de la matriz de transición depende de las probabilidades relativas del estado *m* inicial y del estado *n* final. Existen dos casos a considerar: i) Si $\delta V_{nm} = V_n - V_m \leq 0$ entonces $\rho_n \geq \rho_m$ y se aplica la ecuación (96). ii) Si $\delta V_{nm} > 0$ entonces $\rho_n < \rho_m$ y se aplica la ecuación (97). Nótese que $V_m \equiv V(\Gamma_m)$. En un movimiento MC, el siguiente paso es para determinar δV_{nm}. La determinación de δV_{nm} no requiere un recálculo completo de la energía configuracional del *m*-ésimo estado, solo los cambios del átomo que se mueve. Por ejemplo, en la figura 53 se ilustra el cambio en energía potencial.

Figura 53. *Nueva posición del átomo*



Fuente: elaboración propia.

Este cambio se calcula computando la energía del átomo *i* con todos los demás átomos antes y después del movimiento

$$\delta V_{nm} = \left(\sum_{j=1}^{N} v\left(r_{ij}^{n}\right) - \sum_{j=1}^{N} v\left(r_{ij}^{m}\right) \right)$$

donde la suma sobre los átomos excluye el átomo *i*. Al calcular el cambio de energía, se considera la interacción explícita del átomo *i* con todos sus vecinos afuera de una distancia r_c de corte. La contribución de átomos más allá del corte se podría estimar usando una corrección de campo medio, pero la corrección para el átomo *i* en las posiciones viejas y nuevas es exactamente la misma, y no se necesita incluir explícitamente en el cálculo de δV_{nm}.

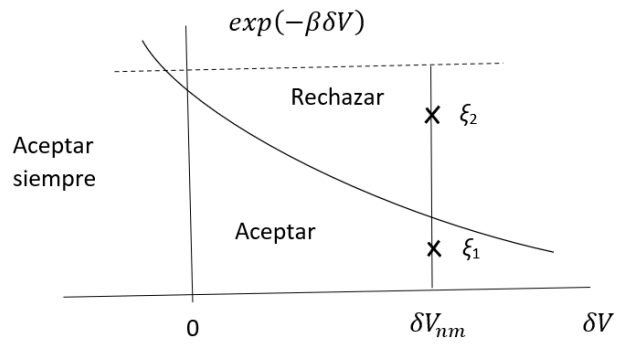
Si el movimiento es hacia abajo en energía ($\delta V_{nm} \leq 0$), entonces la probabilidad del estado *n* es mayor que la del estado *m* y se acepta la nueva configuración. El método de elegir movimientos de prueba asegura que la probabilidad de transición $\pi_{mn} = \alpha_{mn}$ es el valor requerido por la ecuación (96-98).

Si el movimiento es hacia arriba en energía ($\delta V_{nm} > 0$), entonces se acepta el movimiento con una probabilidad ρ_n/ρ_m de acuerdo con la ecuación (97). Otra vez se incluye automáticamente el factor α_{mn} al hacer el movimiento. Esta razón se puede expresar como el factor de Boltzmann de la diferencia de energía:

$$\frac{\rho_n}{\rho_m} = \frac{Z_{NVT}^{-1} exp(-\beta V_n)}{Z_{NVT}^{-1} exp(-\beta V_m)} = \frac{exp(-\beta V_n) exp(-\beta \delta V_{nm})}{exp(-\beta V_n)} = exp(-\beta \delta V_{nm})$$

Al aceptar un movimiento con una probabilidad $exp(-\beta \delta V_{nm})$ se genera un número aleatorio uniformemente en (0,1). El número aleatorio se compara con $exp(-\beta \delta V_{nm})$. Si es menor que $exp(-\beta \delta V_{nm})$ se acepta el movimiento. Este procedimiento se ve en la gráfica de la figura 54.

Figura 54. *Procedimiento de aceptación o rechazo de la nueva posición*



Fuente: Allen y Tildesley (1987).

Durante la corrida, supóngase que se acepta un movimiento particular hacia arriba δV_{nm}. Si en este punto se elige un número ξ_1 aleatorio, el movimiento se acepta. Si se elige ξ_2 el movimiento se rechaza. En el curso de la corrida el resultado neto es que la energía cambia tal que δV_{nm} se acepta con una probabilidad $exp(-\beta\delta V_{nm})$. Si el movimiento hacia arriba se rechaza, el sistema permanece en el estado *m* de acuerdo con la probabilidad finita π_{mm} de la ecuación (98). En este caso, el átomo se retiene en su posición vieja, y la configuración vieja se recuenta como un nuevo estado en la cadena.

Un algoritmo de Metropolis se puede resumir como sigue:

Dado un estado *i* con energía E_i,

Genera un nuevo estado *j* mediante un mecanismo de perturbación

(pequeña distorsión del estado *i*)

Calcula la energía del nuevo estado E_j.

si $\left(E_j - E_i\right) \leq 0$

entonces acepta el estado *j* como estado nuevo

si no acepta el estado con probabilidad: $exp\left(\frac{E_i - E_j}{k_B T}\right)$

Montecarlo isotérmico-isobárico

Una ventaja del método MC es que se puede adaptar fácilmente al cálculo de promedios en cualquier ensamble, y que el número de moléculas, la temperatura y la presión son fijas en el ensamble isotérmico-isobárico, mientras que al volumen de la caja de simulación se le permite fluctuar. Las simulaciones originales de NPT-constante se realizaron en esferas y

discos duros, pero la técnica se extendió para cubrir potenciales continuos en su estudio de mezclas de Lennard-Jones. Este ensamble se pensó apropiado para simular mezclas, dado que las mediciones experimentales de las propiedades en exceso se registran a presión constante, y las teorías de mezclado a menudo se formulan bajo este supuesto.

El método también se ha utilizado en la simulación de fluidos de un solo componente, y en el estudio de las transiciones de fase. Vale la pena recordar que a N, P, T constantes no deberíamos ver dos fases coexistiendo en la misma celda de simulación, un problema que altera la simulación de las transiciones de fase en el ensamble canónico.

En el conjunto de NPT constante, el promedio de configuración de una propiedad viene dado por

$$\langle A \rangle_{NPT} = \frac{\int_0^\infty dV exp(-\beta PV)\, V^N \int d\boldsymbol{s}\, A(\boldsymbol{s}) exp\big(-\beta V(\boldsymbol{s})\big)}{Z_{NPT}}$$

En esta ecuación, Z_{NPT} es la ecuación integral configuracional y V es el volumen del fluido. Nótese que en esta ecuación usamos un conjunto de coordenadas escaladas $\boldsymbol{s} = (\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N)$ donde

$$\boldsymbol{s} = L^{-1}\boldsymbol{r}$$

En este caso, la integral de configuración en la ecuación $\langle A \rangle_{NPT}$ está sobre el cubo unitario y el factor adicional de V^N proviene del elemento de volumen d$\boldsymbol{r}$. (En esta sección se supone que la caja de simulación es un cubo del lado $L = V^{-1/3}$; los argumentos pueden extenderse fácilmente a cajas no-cúbicas).

El esquema de Metropolis se implementa generando una cadena de estados de Markov que tiene una distribución limitante proporcional a

$$exp\big(-\beta\big(PV + V(\boldsymbol{s})\big) + NlnV\big)$$

y el método utilizado es una extensión directa de las ideas discutidas en lo anterior.

Se genera un nuevo estado por desplazamiento de una molécula aleatoriamente y/o haciendo un cambio de volumen aleatorio de V_m a V_n

$$\boldsymbol{s}_i^n = \boldsymbol{s}_i^m + \delta s_{max}(2\boldsymbol{\xi} - \mathbf{1})$$

$$V_{\boldsymbol{n}} = V_{\boldsymbol{m}} + \delta V_{max}(2\xi - 1)$$

donde ξ es un número aleatorio generado uniformemente en (0,1), mientras que $\boldsymbol{\xi}$ es un vector cuyas componentes también son números aleatorios uniformes en (0,1), y $\mathbf{1}$ es el vector (1,1,1). δs_{max} y δV_{max} gobiernan el cambio máximo en las coordenadas escaladas de

las partículas, y en el volumen de la caja de simulación, respectivamente. Sus valores precisos dependerán del punto del estado estudiado, y se eligen para producir una tasa de aceptación de 35-50 por ciento. Estos valores son conjeturas iniciales y se pueden ajustar automáticamente por el programa, aunque en este caso hay dos desplazamientos máximos independientes y muchas combinaciones diferentes producirán una tasa de aceptación dada.

Una vez que se ha producido el nuevo estado *n*, se calcula la cantidad δH,

$$\delta H_{nm} = \delta V_{nm} + P(V_n - V_m) - N\beta^{-1} ln(V_n/V_m)$$

δH_{nm} está estrechamente relacionado con el cambio de entalpía al pasar del estado *m* al estado *n*. Los movimientos se aceptan con una probabilidad $min\big(1, exp(-\beta \delta H_{nm})\big)$ usando las técnicas discutidas en la sección anterior. Un movimiento puede proceder con un cambio en la posición de las partículas o un cambio en el volumen o una combinación de ambos. Podría ser más conveniente hacer cambios aleatorios en *ln V* en lugar de en *V* misma. Se elige un número aleatorio $\delta(ln\, V)$ uniformemente en algún rango $(-\delta(ln\, V)_{max}, \delta(ln\, V)_{max})$, el volumen multiplicado por $exp\big(\delta(ln\, V)\big)$ y las posiciones moleculares escaladas acordadamente. El único cambio en el procedimiento de aceptación / rechazo es que el factor N en la ecuación anterior se reemplaza por N + 1.

Una diferencia importante entre este ensamble y el ensamble canónico es que cuando un movimiento implica un cambio en el volumen, la densidad del líquido cambia. En este caso, las correcciones de largo-alcance a la energía en los estados *m* y *n* son diferentes y deben incluirse directamente en el cálculo de δV_{nm}.

En el caso general, cambiar el volumen es computacionalmente más costoso que desplazar una molécula. Para el desplazamiento de una molécula, a lo sumo hay $2(N-1)$ cálculos del potencial par al calcular δV_{nm}. En general, un cambio de volumen en un fluido par-aditivo requiere el recálculo de todas las $\frac{1}{2}N(N-1)$ interacciones. Afortunadamente, para los potenciales más simples, el cambio en *V* con el volumen se puede calcular escalando. Como ejemplo, considere la energía configuracional de un fluido Lennard-Jones en el estado *m*:

$$V_m = 4\varepsilon \sum_i \sum_{j>i} \left(\frac{\sigma}{L_m s_{ij}^m}\right)^{12} - 4\varepsilon \sum_i \sum_{j>i} \left(\frac{\sigma}{L_m s_{ij}^m}\right)^{6}$$
$$= V_m^{(12)} + V_m^{(6)}$$

Aquí hemos dividido el potencial en sus componentes separados de doceava y sexta potencia. Si el único cambio entre los estados *m* y *n* es la longitud de la caja, entonces la energía del nuevo estado es

$$V_n = V_m^{(12)} \left(\frac{L_m}{L_n}\right)^{12} + V_m^{(6)} \left(\frac{L_m}{L_n}\right)^{6}$$

y

$$\delta V_{nm} = \delta V_{nm}^{\text{vol}} = V_m^{(12)} \left[\left(\frac{L_m}{L_n}\right)^{12} - 1\right] + V_m^{(6)} \left[\left(\frac{L_m}{L_n}\right)^{6} - 1\right]$$

Este cálculo es extremadamente rápido y solo requiere que los dos componentes de la energía potencial, $V^{(12)}$ y $V^{(6)}$ se almacenen por separado. Si el límite potencial de corte se lleva a escala con la longitud de la caja (es decir, $r_c = s_c L$ con s_c constante), entonces los términos separados $V_{LRC}^{(12)}$ y $V_{LRC}^{(6)}$ escalan tal como $V^{(12)}$ y $V^{(6)}$, respectivamente. Si además de a un cambio de longitud de caja, una molécula se desplaza simultáneamente, entonces hay dos contribuciones

$$\delta V_{nm} = \delta V_{nm}^{dis} + \delta V_{nm}^{vol}$$

donde δV_{nm}^{vol} está dado por la ecuación previa a la anterior, y

$$\delta V_{nm}^{dis} = V_n(L_n) - V_m(L_n)$$

Así, el cambio de energía en el desplazamiento se obtiene usando la nueva longitud-de-caja L_n (piense en escalar la caja, seguido de mover la molécula).

Este simple procedimiento para el cálculo de δV_{nm} se basa en que solo hay una longitud característica en la función potencial. Este puede no ser el caso para algunos potenciales pares complicados, y tampoco es cierto para la mayoría de los modelos moleculares, donde aparecen longitudes de enlace intramoleculares, así como potenciales sitio a sitio. Para un modelo de sitio de interacción, el simple escalamiento implicaría un cambio no-físico en la forma molecular. Para estos casos, el cálculo de δV_{nm}^{vol} es costoso y, por lo tanto, los cambios de volumen deben realizarse mucho menos frecuentemente que el desplazamiento de una partícula.

Al promediar sobre los estados en la cadena de Markov, es posible calcular propiedades mecánicas como el volumen y la entalpía, y varias propiedades relacionadas con sus fluctuaciones. En común con la simulación de NVT constante, este método solo muestrea regiones importantes del espacio de fase y no es posible calcular propiedades *estadísticas*

como la energía libre de Gibbs. Durante el curso de una corrida particular, el virial se puede calcular de la manera habitual para producir una estimación de la presión. Esta presión calculada (incluyendo la corrección de largo-alcance) debería ser igual a la presión de entrada, P, utilizada en la ecuación para δH_{nm} que genera la cadena de Markov. Esta prueba es una comprobación útil de un programa a NPT-constante apropiadamente codificado correctamente.

Finalmente, se realizan simulaciones a presión constante de discos y esferas duros usando los métodos descritos en esta sección.

Líquidos moleculares

Moléculas rígidas

En la simulación MC de un líquido molecular, la matriz subyacente de la cadena de Markov se altera para permitir movimientos que generalmente consisten en una translación y rotación combinadas de una molécula. Cadenas que involucran un número de pasos ya sean puramente traslacionales o rotacionales son perfectamente apropiadas, pero generalmente no se explotan en la simulación de líquidos moleculares. (Ha habido simulaciones de modelos idealizados de cristales líquidos y cristales plásticos donde los centros de las moléculas están fijos a una red tridimensional. Estas simulaciones consisten en movimientos puramente rotacionales. La parte traslacional del movimiento se lleva a cabo desplazando aleatoriamente el centro de masa de una molécula a lo largo de cada uno de los ejes fijos-al-espacio. Como antes, el desplazamiento máximo se rige por el parámetro ajustable δr_{max}. La orientación de una molécula a menudo se describe en términos de los ángulos de Euler. Se puede lograr un cambio de orientación tomando pequeños desplazamientos aleatorios en cada uno de los ángulos de Euler de la molécula i.

$$\phi_i^n = \phi_i^m + (2\xi_1 - 1)\delta\phi_{max}$$
$$\theta_i^n = \theta_i^m + (2\xi_2 - 1)\delta\theta_{max}$$
$$\psi_i^n = \psi_i^m + (2\xi_3 - 1)\delta\psi_{max}$$

Donde $\delta\phi_{max}, \delta\theta_{max}, y\ \delta\psi_{max}$ son los desplazamientos en los ángulos de Euler.

En un paso MC la tasa de probabilidades de los dos estados está dada por

$$\frac{\rho_n}{\rho_m} = \frac{exp\big(-\beta(V_m + \delta V_{nm})\big)d\boldsymbol{r}^n d\Omega^n}{exp(-\beta V_m)d\boldsymbol{r}^m d\Omega^m}$$

Se han incluido los elementos de volumen apropiados para convertir las densidades de probabilidad en probabilidades. $d\Omega^m = \prod_{i=1}^{N} \mathrm{d}\Omega_i^m$ y $d\Omega^m = sen\theta_i^m d\theta_i^m d\psi_i^m d\phi_i^m/\Omega$ para la molécula *i* en el estado *m*. La constante $\Omega = 8\pi^2$ para moléculas no-lineales. En el caso de moléculas lineales Ω no se requiere para definir la orientación, y $\Omega = 4\pi$. Los elementos de volumen para los estados *m* y *n* no han sido incluidos previamente en la tasa ρ_n/ρ_m por la simple razón de que son los mismos en ambos estados para un movimiento traslacional y se cancelan. Para un movimiento que solo involucra la molécula *i*

$$\frac{\rho_n}{\rho_m} = exp(-\beta\delta V_{nm})\frac{sen\theta_i^n}{sen\theta_i^m}$$

La tasa de los senos debe aparecer en la matriz de transición π_{nm} ya sea en el criterio de aceptación/rechazo o en el elemento de matriz π_{nm} subyacente. Esta última aproximación es más conveniente. Equivale a elegir desplazamientos aleatorios en $cos\theta_i$ en lugar de en θ_i:

$$cos\theta_i^n = \cos\theta_i^m + (2\xi_2 - 1)\delta(cos\theta)_{max}$$

y adoptando la propuesta habitual de Metropolis de aceptar o rechazar con una probabilidad de $min\big(1, exp(-\beta\delta V_{nm})\big)$. La inclusión del factor $sen\theta$ en la cadena subyacente evita dificultades con $\theta_i^m = 0$. Las ecuaciones para ϕ_i^n, ψ_i^n y $cos\theta_i^n$ mueven una molécula de un estado de orientación a cualquiera de sus estados de orientación vecinos con igual probabilidad, y cumplen la condición de reversibilidad microscópica.

Es útil mantener los ángulos que describen la orientación de una molécula particular en el rango apropiado (- π, π) para ψ y ϕ, y (0, π) para θ. Esto no es esencial, pero evita trabajo innecesario y posible desbordamiento en la evaluación posterior de cualquier función trigonométrica.

Un método alternativo para rotar las moléculas en una simulación MC en agua implica seleccionar una molécula y rotarla una cantidad aleatoria $\delta\gamma$ alrededor de uno de los tres ejes fijos-en-el-espacio elegidos al azar. Por ejemplo, en el caso de un fluido de moléculas lineales es más conveniente representar la orientación molecular mediante un vector unitario $\hat{\boldsymbol{e}}$ fijo en la molécula. La orientación de la molécula *i* se representa por un vector con componentes

$$e_{ix} = cos\phi_i sen\theta_i$$
$$e_{iy} = sen\phi_i sen\theta_i$$
$$e_{iz} = cos\theta_i$$

Se genera una nueva configuración usando

$$\boldsymbol{e}_i^n = \boldsymbol{A}_x \boldsymbol{e}_i^m$$

donde

$$\boldsymbol{A}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\delta\gamma & sen\delta\gamma \\ 0 & -sen\delta\gamma & cos\delta\gamma \end{pmatrix}$$

y corresponde a una rotación de $\delta\gamma$ alrededor del eje x fijo-en-el-espacio. Existen ecuaciones similares para rotaciones alrededor tanto del eje y como del eje z,

$$\boldsymbol{A}_y = \begin{pmatrix} cos\delta\gamma & 0 & -sen\delta\gamma \\ 0 & 1 & 0 \\ sen\delta\gamma & 0 & cos\delta\gamma \end{pmatrix}$$

y

$$\boldsymbol{A}_z = \begin{pmatrix} cos\delta\gamma & sen\delta\gamma & 0 \\ -sen\delta\gamma & cos\delta\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La ventaja de este método es que la orientación de la molécula se puede almacenar como uno o más vectores. Para fluidos del modelo de sitio de interacción, esto significa que la evaluación costosa de las funciones trigonométricas se puede evitar por completo. Este no es el caso con los métodos que implican cambios en los ángulos de Euler. Fórmulas similares se aplican cuando los cuaterniones se usan para representar orientaciones moleculares.

En un tercer método para cambiar la orientación de una molécula se elige la nueva orientación de prueba, $\boldsymbol{e}_i^n$, aleatoria y uniformemente en una región de la superficie de una esfera con la restricción de que

$$1 - \boldsymbol{e}_i^n \cdot \boldsymbol{e}_i^m < d \ll 1$$

d controla el tamaño del desplazamiento máximo y una primera aproximación sensata para este parámetro es 0.2. Existen otros métodos para generar un vector aleatorio en la superficie de una esfera. Estos pueden combinarse fácilmente con el método para generar aleatoria y uniformemente en una región restringida para producir un algoritmo simple que generará orientaciones con la restricción de la ecuación anterior.

Una dificultad con los métodos de MC para fluidos moleculares es que generalmente hay una serie de parámetros que gobiernan el desplazamiento máximo de traslación y orientación de una molécula durante un movimiento. Como de costumbre, estos parámetros se pueden ajustar automáticamente para dar una tasa de aceptación de ≈0.5, pero no existe

un conjunto único de parámetros de desplazamiento máximo que logre esto. Un conjunto sensible de valores se obtiene mejor mediante ensayo y error para la simulación particular en turno.

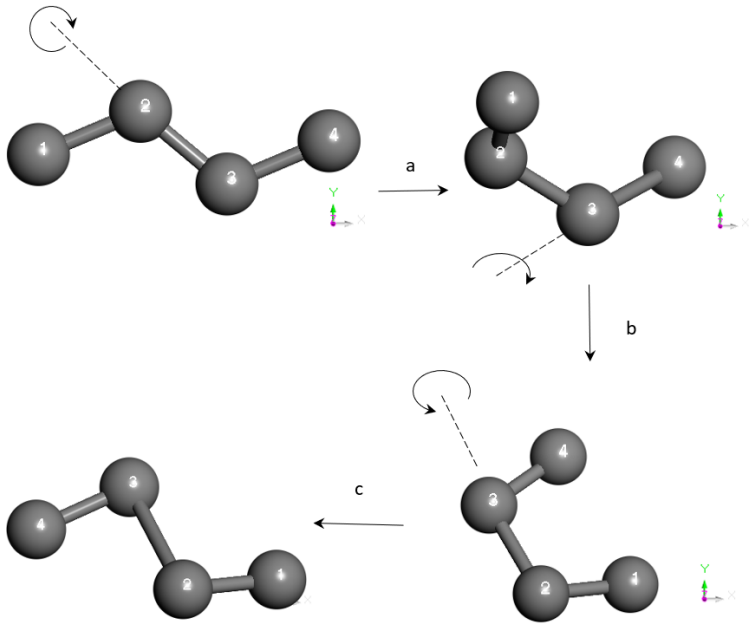
Moléculas no-rígidas

La no rigidez introduce nuevas dificultades en la técnica de MC. El problema en este caso es encontrar un conjunto adecuado de coordenadas generalizadas para describir las posiciones y momentos de las moléculas. Una vez que se han establecido las coordenadas de momento generalizadas, las integraciones sobre los momentos pueden realizarse analíticamente, lo que dejará solo la parte configuracional del ensamble promedio. Sin embargo, la integración sobre los momentos producirá jacobianos complicados en la integral configuracional, uno para cada molécula. El jacobiano será una función de las variables orientativas generalizadas, θ, ϕ que describen la orientación general de la molécula y los ángulos de flexión y torsión del enlace que describen la configuración interna. Un ejemplo simple de este tipo de término es el $sen\theta_i$ en la integral configuracional para moléculas rígidas, que proviene de la integración sobre el momento $\left(p_\phi\right)_i$. Estos jacobianos son importantes para calcular la relación ρ_n/ρ_m utilizada para generar la cadena de Markov en el método de Metropolis o, correspondientemente, para el diseño correcta de la matriz estocástica subyacente. Para las moléculas no-rígidas, el manejo correcto de los términos jacobianos es más difícil.

Este problema puede resolverse satisfactoriamente para la clase de moléculas no-rígidas donde el momento total de inercia es independiente de las coordenadas de rotación interna (por ejemplo iso-butano, acetona). Las coordenadas generalizadas también se han desarrollado para un modelo no-rígido de butano, que no cae en esta clase simple, pero las expresiones son complicadas y lo son cada vez más para moléculas más largas.

Una forma de trabajar con coordenadas generalizadas es la siguiente. En butano es posible restringir las longitudes de enlace y los ángulos de flexión de enlace, al tiempo que permite que el ángulo de torsión cambie de acuerdo con su función potencial. El movimiento de la molécula en la simulación se logra mediante movimientos al azar de átomos elegidos aleatoriamente sujetos a las restricciones requeridas. Un ejemplo de tal técnica se muestra para el butano en la figura 55

Figura 55. *Método para mover una cadena sujeta a las constricciones de longitud y ángulo de enlace*



Fuente: elaboración propia.

Una secuencia típica de MC podría ser:

(a) el átomo 1 se mueve girando alrededor del enlace 2-3;

(b) los átomos 1 y 2 se mueven simultáneamente girando alrededor del enlace 3-4; y

(c) el átomo 4 se mueve girando alrededor del enlace 2-3.

Los movimientos (a) y (c) implican un desplazamiento aleatorio del ángulo de torsión ϕ, en el intervalo $(-\pi,\pi)$. La molécula completa se traslada y rota a través del espacio haciendo rotaciones aleatorias de átomos alrededor de enlaces elegidos al azar. También podemos incluir una translación explícita de la molécula completa y una rotación total alrededor de uno de los ejes fijos en el espacio. La desventaja de esta simple aproximación de alta densidad es que una pequeña rotación alrededor del enlace 1-2 puede causar un movimiento sustancial del átomo 4, lo que probablemente resulte en traslape y una tasa alta de rechazo para nuevas configuraciones.

Si consideramos el caso de la molécula de butano simplificada introducida en el capítulo de dinámica molecular, entonces un movimiento MC de prueba podría consistir en una traslación y una rotación de la molécula completa, y un cambio en la configuración

interna realizada al elegir un incremento aleatorio en d(cosθ), d(cosθ') y dϕ (ver figura 55). Para evitar los artilugios asociados con la aproximación de restricción, la cadena de Markov debería generarse con una distribución limitante proporcional a

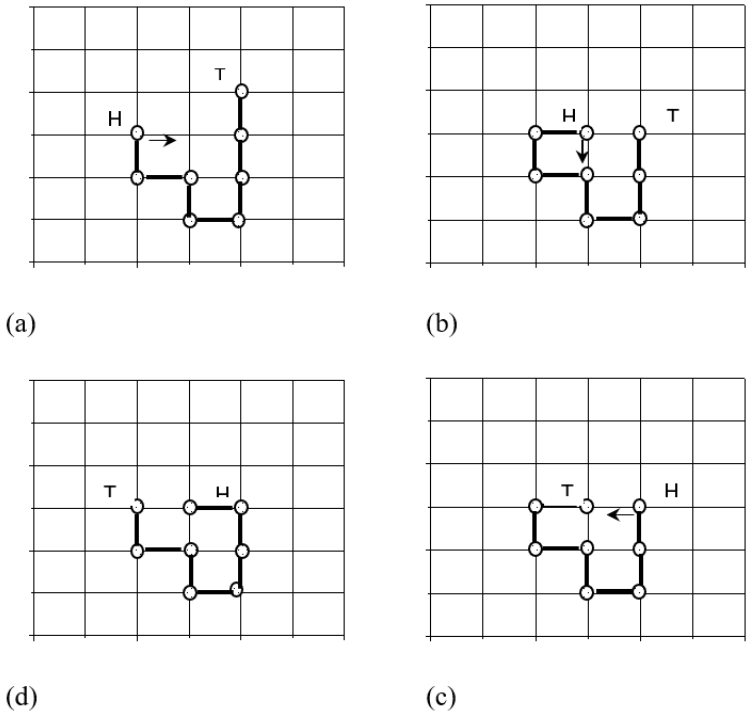
$$exp\big(-\beta(V - V_c)\big) = exp\left(-\beta\left(V + \frac{1}{2}k_b T ln[2 + sen^2\theta + sen^2\theta']\right)\right)$$

Si θ y θ' permanecen cerca de sus valores de equilibrio, podría ser posible introducir solo un pequeño error al despreciar el potencial de restricción V_c en la ecuación anterior. El término de restricción se vuelve más complicado e importante en el caso de restricciones de ángulo de enlace. Por esta razón, ha habido pocas simulaciones de Metropolis MC de moléculas flexibles de cadena larga. La técnica de elección es la dinámica de restricción, utilizando potenciales de ángulo de enlace cuadrático. Las cadenas individuales se pueden simular utilizando métodos crudos de MC. En esta técnica, una cadena de polímero de longitud especificada se construye aleatoriamente en el espacio o en una red. Una cadena se abandona si se introduce un traslape sustancial durante su construcción. Cuando se ha producido un gran número N de cadenas de la longitud requerida, el promedio de una propiedad (como la distancia de extremo a extremo) se calcula a partir de

$$\langle A \rangle = \frac{\sum_{i=1}^{N} A_i exp(-\beta V_i)}{\sum_{i=1}^{N} exp(-\beta V_i)}$$

donde las sumas se extienden sobre todas las cadenas de N polímeros. El enfoque no es aplicable para un fluido denso de cadenas. En un fluido real, es probable que una cadena se mueva de manera deslizante: la cabeza de la cadena se mueve a una nueva posición y el resto de la cadena sigue como una serpiente o lagarto. Este tipo de movimiento se denomina *reptación*. Un algoritmo MC exitoso imitaría este movimiento. El modelo de *serpiente deslizante* se aplicó originalmente a un polímero en una red bidimensional y se muestra un ejemplo simple en la figura 56.

Figura 56. *Movimiento de reptación*



Nota. Las configuraciones son generadas en el orden (a), (b), (c) y (d).

Fuente: Allen y Tildesley (1987).

Se simula un polímero de ocho segmentos. Un extremo de la molécula se elige al azar para ser la cabeza (H) mientras que el otro es la cola (T). La cabeza se mueve a una nueva posición en la red, los demás átomos se mueven un sitio a lo largo de la cadena y la posición de la cola queda vacante. En la figura 56b se rechaza el movimiento de cabeza propuesto, ya que no se permite que las cadenas se superpongan. La configuración resultante, que es idéntica a la anterior, se incluye en el promedio y la simulación procede. Podemos ver que la selección aleatoria de la cabeza y la cola es importante, ya que de lo contrario el sistema podría bloquearse con la cabeza incapaz de moverse. En la figura 56c, la cabeza y la cola se han intercambiado, y se acepta el movimiento propuesto, ya que la posición de la cola de la cadena estará vacía cuando se complete el movimiento. Aunque este ejemplo ilustra el método para una sola cadena de polímero, se extiende fácilmente a un fluido denso de cadenas, ya que solo un átomo se mueve al generar cada nueva configuración.

Se han aplicado algoritmos de *reptation MC* a cadenas en una red tridimensional y a fluidos continuos, con potenciales arbitrarios intermoleculares e intramoleculares. El método

explota la solución Metropolis a la matriz de transición para muestrear asintóticamente la distribución de Boltzmann. En el caso del fluido, todos los átomos en el interactúan a través de la parte repulsiva del potencial de Lennard-Jones. Esta interacción controla el volumen excluido de las cadenas. Además, los átomos adyacentes en la misma cadena interactúan a través de un potencial armónico modificado:

$$v^H(r) = \begin{cases} -0.5k\sigma_1^2 ln[1-(r/\sigma_1)^2] & 0 \le r \le \sigma_1 \\ \infty & r > \sigma_1 \end{cases}$$

donde, típicamente, $\sigma_1 = 1.95\sigma$ y $k = 20$. Cada cadena se considera a su vez, y se elige un extremo al azar como la cabeza. Las coordenadas iniciales de los átomos en la *i*-ésima cadena son $(\boldsymbol{r}_{i1}, \boldsymbol{r}_{i2}, \dots, \boldsymbol{r}_{in_a})$. Se selecciona una nueva posición para la cabeza, átomo n_a,

$$\boldsymbol{r}' = \boldsymbol{r}_{in_a} + \delta\boldsymbol{r}$$

La dirección de $\delta\boldsymbol{r}$ se elige al azar en la superficie de una esfera, y la magnitud δr se elige de acuerdo con la distribución de probabilidad $exp(-\beta v^H(r))$ utilizando una técnica de rechazo (véase el apartado **Importancia del muestreo**). Por lo tanto, el potencial de enlace intramolecular se utiliza para seleccionar el movimiento de prueba. La cadena tiene una nueva configuración de prueba $(\mathbf{r}_{i1}, \mathbf{r}_{i2}, \dots, \mathbf{r}_{in_a}, \mathbf{r}')$. El cambio en las interacciones no-enlazadas al crear una nueva configuración se calcula sumando sobre todos los átomos.

$$\delta V_{nm} = \sum_{a=2}^{n_a} v^{RLJ}(|\mathbf{r}_{ia} - \mathbf{r}'|) - v^{RLJ}(|\mathbf{r}_{ia} - \mathbf{r}_{i1}|) + \sum_{j\neq i}\sum_{a=2}^{n_a} v^{RLJ}(|\mathbf{r}_{ja} - \mathbf{r}'|) - v^{RLJ}(|\mathbf{r}_{ja} - \mathbf{r}_{i1}|)$$

Aquí se incluyen las interacciones no-enlazantes de la cadena *i* y de las demás cadenas *j*. δV_{nm} se usa para decidir si el movimiento debe ser aceptado o rechazado de acuerdo con los criterios de Metropolis. Como de costumbre, se cuentan los movimientos rechazados. El enfoque funciona bien; no hay restricciones geométricas para tener en cuenta en este ejemplo, ya que todos los átomos pueden moverse libremente bajo la influencia de los potenciales.

Capítulo 5. Uso de software de simulación molecular

La simulación molecular se describe mejor a partir de dos modelos básicos: i) dinámica molecular (DM) y ii) Montecarlo (MC). El paquete más robusto para calcular DM es GROMACS, entre muchos otros paquetes libres. Para tener un buen funcionamiento de la DM se requiere de potenciales de interacción, y aunque Lennard-Jones es el más empleado, la dinámica molecular cuántica usa directamente mecánica cuántica para proporcionar una mejor manera de calcular dichos potenciales. BIOVIA Materials Studio tiene módulos para efectuar automáticamente dichos cálculos. No obstante, existe un gran número de paquetes libres para llevar a cabo cálculos cuánticos *ab-initio*, como: HONDO, PSHONDO-IJKL-GMCP-FOK-CIPSI, GAMESS, NWChem, Gaussian (aunque las versiones más actuales son de Licencia), etcétera.

GROMACS es un motor para realizar minimización de energía y simulaciones de dinámica molecular. El modelado molecular proporciona un proceso para describir sistemas químicos complejos en términos de un modelo atómico realista, con el objetivo de comprender y predecir las propiedades macroscópicas basadas en el conocimiento detallado a escala atómica. El modelado molecular se utiliza entre otras cosas para diseñar nuevos materiales, que requieren una predicción precisa de propiedades físicas de sistemas realistas. La dinámica molecular es una técnica de simulación molecular computacional que permite analizar el comportamiento o evolución de un sistema físico, químico o biológico a través del tiempo, calculando las fuerzas entre los átomos que lo conforman, mediante las ecuaciones de movimiento de Newton, generando trayectorias de un sistema compuesto de N partículas con alta resolución temporal y espacial.

Una ventaja de este tipo de simulación es que, por medio de un visualizador, VMD, por ejemplo, se pueden ver las estructuras en movimiento. A partir de esta dinámica, se obtienen valores de diferentes propiedades macroscópicas tanto estáticas como dinámicas, tales como energía libre, función de distribución radial, factor de estructura, entropía, solubilidad, viscosidad, presión, temperaturas, envolventes de fase, etc. Estas simulaciones se utilizan en el estudio de proteínas y biomoléculas, así como en ciencia de materiales.

Las diferentes etapas de una simulación por dinámica molecular (DM) se pueden resumir en los siguientes puntos:

a) Preparación del sistema: se establece o elige la configuración inicial a utilizar conformada por N átomos, y se definen las condiciones de equilibrio.
b) Equilibramiento: mediante diferentes esquemas o ensambles (microcanónico, canónico, isotérmico-isobárico) se lleva a cabo el equilibrio de la simulación, considerando que la energía permanece constante. Dependiendo del esquema a utilizar se pueden definir condiciones de energía, presión y/o temperatura.
c) Producción: una simulación de DM se define como el algoritmo que realiza el cálculo de las fuerzas, posiciones y velocidades del sistema, con lo que se pueden obtener propiedades fisicoquímicas, tiempo y pasos de ejecución.
d) Resultados: finalmente se analizan los resultados y se obtiene información, aprovechando la trayectoria y el comportamiento del sistema por medio de visualizadores moleculares con interfaz gráfica.

Metodología de una dinámica molecular

Tanto la metodología como los algoritmos de simulaciones de DM computacionales se han construido poco a poco, día con día. Para realizar una dinámica molecular se requiere una configuración inicial del sistema, es decir, se requieren las posiciones iníciales de los átomos que lo conforman, su topología y las condiciones iníciales del sistema. Los pasos principales en una simulación de DM son los siguientes:

1) Especificar el sistema de átomos o moléculas, lo cual incluye el modelo de las interacciones moleculares o atómicas, el modelo de las interacciones ambiente-sistema, y el desarrollo de ecuaciones de movimiento.
2) Resolver las ecuaciones de movimiento para generar las trayectorias moleculares o atómicas.
3) Analizar las trayectorias para obtener las propiedades de interés. Las interacciones moleculares se representan a través de la energía potencial. Una aproximación muy utilizada consiste en descomponer la energía potencial en términos que involucran las interacciones de pares o tripletes de átomos. Estos potenciales comúnmente parten de una formulación empírica, tal como el potencial de Lennard-Jones, desarrollado para el análisis del comportamiento a nivel atómico en gases nobles.

Otros enfoques para calcular las formas de los potenciales interatómicos de diferentes tipos de átomos provienen de los primeros principios, partiendo de simplificaciones de la ecuación de Schrödinger para obtener soluciones aproximadas, ya que ésta no puede resolverse de manera precisa para sistemas de muchos átomos. Las interacciones ambiente-sistema involucran, en parte, lo relacionado con las condiciones de frontera. La selección de estas condiciones depende de la situación física que se desea simular en un sistema infinito. Por ejemplo, para estudiar el comportamiento de las partículas cercanas de carbono con pirrol se pueden usar condiciones de integración de las ecuaciones de movimiento. Esto se hace a través de algoritmos como el de Verlet (o su variante, el método *leap frog*), y el predictor-corrector de Gear, entre otros.

Los algoritmos anteriores para resolver las ecuaciones diferenciales aplican métodos de diferencias finitas para la integración de las ecuaciones diferenciales, y se realiza en etapas de tiempo Δt muy pequeñas. Para cada Δt se realiza el siguiente procedimiento:

i) Se calcula la energía potencial asociada a una conformación de la estructura. Para ello se utiliza un funcional clásico muy sencillo denominado campo de fuerzas (*force-field*) que ha sido parametrizado para reproducir resultados experimentales en sistemas modelo, o cálculos mecánico-cuánticos de alto nivel.

ii) Se obtiene la derivada de la energía respecto a las posiciones atómicas, la cual nos permite obtener la fuerza que actúa sobre cada partícula del sistema.

iii) La segunda ley de Newton permite entonces asignar las aceleraciones que actuarían sobre cada partícula (de masa conocida) de la molécula.

iv) Escogiendo un incremento de tiempo tan pequeño que podamos suponer que en él la aceleración es constante, podemos obtener numéricamente las velocidades que actúan sobre cada partícula del sistema por simple integración de las aceleraciones respectivas.

v) Finalmente, la integración de las nuevas velocidades nos permite obtener las nuevas posiciones asignadas a cada partícula de la estructura

Dado que en una dinámica molecular el número de partículas es una variable relevante, se debe realizar el cálculo considerando el tipo de ensamble. Esta consideración depende del cálculo que se desea llevar a cabo. Los tipos de ensamble usados en la dinámica molecular son los siguientes:

Tipos de ensambles: $\begin{cases} Ensamble\ microcanónico:\ N,V,E\ (exacta) \\ Ensamble\ canónico:\ N,V,T\ (fija) \\ Esamble\ isotérmico-isobárico:\ N,P,T\ (constantes) \end{cases}$

en donde la variable N representa número de partículas, V volumen, E energía, T temperatura y P presión. Nótese que N es común en estos tres ensambles, y es el parámetro que es exacto, fijo o constante. Los ensambles microcanónicos son aislados, los canónicos son sistemas cerrados; sin embargo, si existe transferencia de energía y si hay transferencia de partículas entonces el sistema es abierto y en contacto térmico como es el caso del ensamble gran canónico: μ,V,T. En este ensamble el potencial químico es fijo, y el número de partículas fluctúa. Este último ensamble no se usa directamente y no es de uso común en GROMACS, sin embargo, si existe evidencia de cálculos de equilibrio químico con dicho paquete.

La elección del algoritmo depende de diferentes factores, como esfuerzo computacional en cada paso, tamaño de Δt, conservación de la energía (tipos de ensamble microcanónico), requerimientos de memoria y espacio en disco, precisión. En general, el algoritmo óptimo debe demandar poco costo computacional, debe ser preciso, posibilitar trayectorias largas en tiempos razonables, ser fácil de modificar y adaptarse adecuadamente a diferentes condiciones de simulación.

Pero siempre se debe considerar que la calidad de los resultados (*output*) depende de la información inicial o datos de entrada (*input*).

En cada caso simulado por DM con Gromacs, se señalan los requerimientos técnicos necesarios (procesador, espacio libre en disco duro, memoria RAM, tarjeta gráfica, etc.). Sistemas pequeños compuestos por pocos átomos se simulan en un tiempo corto en computadoras ordinarias, pero sistemas biológicos complejos con ADN o proteínas requieren mayor tiempo de cálculo, desde horas hasta semanas o meses, por esta razón se requiere de equipo con excelente capacidad de cómputo. Aunque sea posible realizar dinámica molecular en cualquier computadora actual, el tiempo que dure la simulación queda en función de las características del sistema a estudiar, y del tiempo que se desee realizar la simulación. La dinámica molecular calcula las fuerzas de los diferentes átomos en función del tiempo; en consecuencia, cada vez que haya un cambio de posición y velocidad, se deben calcular las

propiedades del sistema en cuestión, lo cual exige alta capacidad en infraestructura computacional.

Propiedades macroscópicas

De acuerdo con la mecánica clásica y una vez conocidas las diferentes trayectorias se pueden calcular diferentes propiedades macroscópicas. Sin embargo, debido a la imprecisión en el conocimiento de las trayectorias, algunas propiedades no pueden determinarse con la precisión adecuada. Las propiedades pueden dividirse en:

- Dinámicas: vienen determinadas por la correlación del movimiento a diferentes tiempos. Un ejemplo característico es el coeficiente de difusión (D).
- Estáticas: no dependen del movimiento de las partículas, por ejemplo la capacidad calorífica, el calor de formación, el volumen.

Otra división de las propiedades es la siguiente:

i. Las que dependen de fenómenos colectivos, por ejemplo la constante dieléctrica, la polarizabilidad, etcétera.
ii. Las que dependen del movimiento del tipo de partícula que forma la materia (coeficiente de sedimentación, coeficiente de difusión, etcétera).

La simulación de DM consiste en resolver el sistema de N ecuaciones diferenciales. Una simulación de DM consta de tres etapas:

1) Iniciación: en esta etapa cada partícula se coloca en un lugar determinado del espacio y se la aplica una velocidad al azar, dependiendo de la temperatura que tenga el sistema. Se deja evolucionar el sistema con el tiempo, renormalizando las velocidades de las partículas para que el momento lineal del sistema sea cero y las velocidades sean consecuentes con la temperatura del sistema. El propósito de esto último es relajar el sistema eliminando el exceso de energía (se impone al sistema un control de la temperatura).
2) Continuación de la primera etapa, pero sin renormalizar las velocidades, con el fin de estar seguro de que se ha alcanzado el verdadero equilibrio. Durante este proceso se obtiene y conserva información sobre la historia de las partículas, fundamentalmente sus trayectorias, es decir sus posiciones y velocidades en función del tiempo, ya que a partir de ellas se obtienen diferentes propiedades del sistema

3) Obtención de resultados: se realiza un análisis de las trayectorias para determinar algunas propiedades dinámicas y estáticas del sistema.

Resumiendo, las simulaciones de DM consisten en la integración de las ecuaciones de movimiento de Newton para determinar la manera en que varía la posición de las partículas en función del tiempo. Mediante la adición de una cierta energía cinética especificada mediante la temperatura, las partículas que forman el sistema pueden sobrepasar ciertas barreras de potencial. Estos tipos de cálculos pueden realizarse con dos propósitos diferentes:

- ✓ Determinación de propiedades macroscópicas: el conocimiento de las trayectorias de las partículas del sistema permite la predicción de muchas propiedades macroscópicas del sistema, tales como fluidos (gases reales, líquidos, ...).
- ✓ Salir de mínimos locales: la adición de una energía adicional al sistema puede permitir que se salga de un mínimo local para alcanzar un mínimo global.

Características de GROMACS

GROMACS es un *software* libre, en constante evolución para realizar simulaciones de dinámica molecular, es decir, simular las ecuaciones de Newton del movimiento de sistemas con cientos de millones de partículas.

Es fácil de usar dado que utiliza compiladores en Fortran con topologías y archivos de parámetros escritos en forma de texto, y mostrando mensajes claros de error.

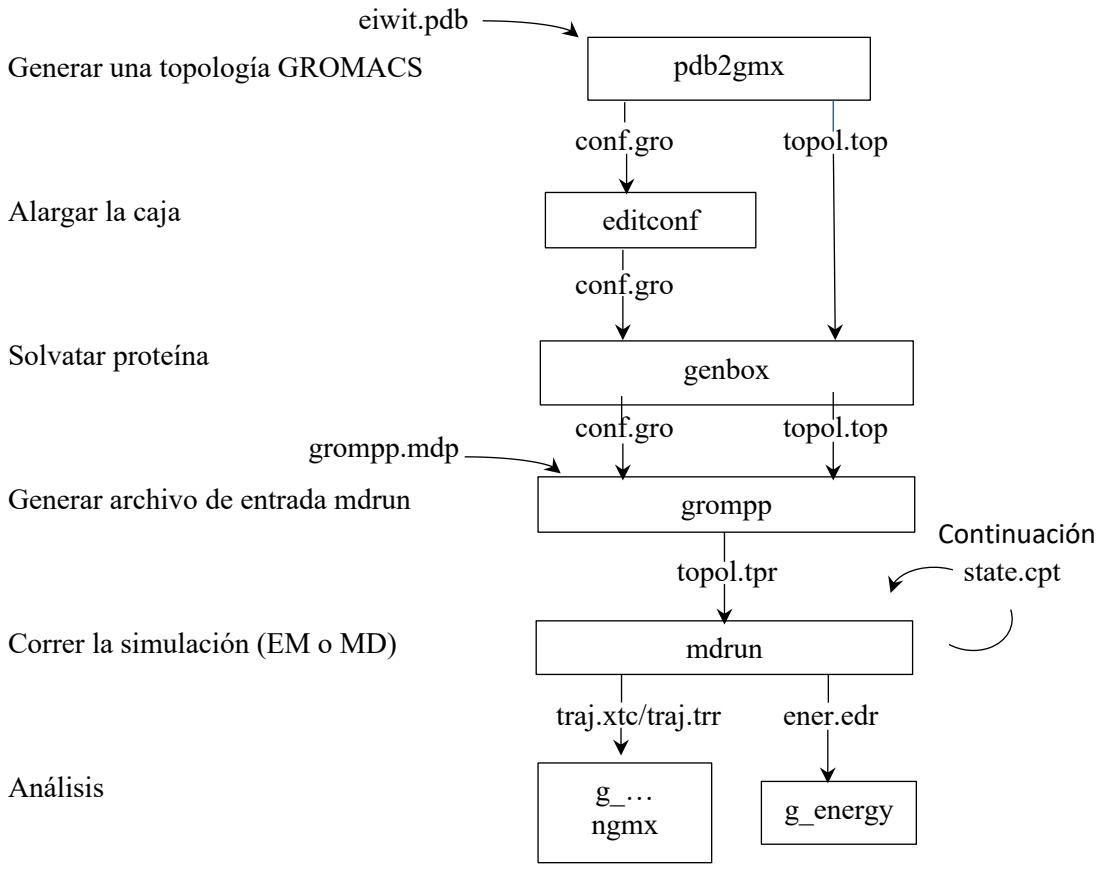
Todos los programas utilizan una interfaz sencilla con opciones de línea de comandos para archivos de entrada y de salida. Siempre se puede obtener ayuda sobre las opciones mediante el uso de la opción −h.

Conforme se lleva a cabo la simulación, GROMACS va informando del punto en que se encuentra la misma, y de la fecha y hora en que espera terminar. Dispone de herramientas flexibles para el análisis de salida de la simulación en pantalla (*ngmx*, *xmgrace*, *g_energy*, etc.)

Generación de archivos para GROMACS

Para simular con GROMACS, previamente se deben generar archivos de entrada con datos de las moléculas que se van a analizar.

Junto con el paquete GROMACS se incluyen programas necesarios para las distintas etapas que conlleva la preparación de los archivos de entrada. En el siguiente diagrama se detallan las etapas típicas de una simulación:



GROMACS puede utilizar como entrada archivos *.pdb* del ProteinDataBank (www.pdb.org), y convertirlos a su formato nativo *.gro* + *.top* con el comando "*pdb2gmx*". Este comando toma el archivo *.pdb* y genera dos archivos, un *.gro* con los átomos y sus coordenadas x-y-z, y un *.top* con las masas atómicas, sus cargas y enlaces.

Una vez que se han creado estos archivos (*.gro* y *.top*), es necesario especificar la "caja" en que alojar los átomos. Esto se hace ejecutando el comando "*editconf*". Este comando toma el archivo *.gro* y lo modifica añadiendo en la última línea las dimensiones de la caja que se le han indicado por parámetros.

Ahora que las dimensiones de la caja se han especificado, es posible agregar más de una de las moléculas del archivo *.pdb* y rellenar el resto de la caja con moléculas disolventes. Ambas acciones son realizadas con el comando "*genbox*". Este comando modifica los archivos *.gro* y *.top* para que contengan todos los átomos que han sido agregados en este paso.

En algunas circunstancias, puede ser necesario introducir iones en la solución, ya sea para neutralizar la carga total en el sistema, para eliminar interacciones electrostáticas entre las moléculas disueltas o para ambas cosas. Estos iones se agregan al sistema mediante la utilidad "*genion*".

Una vez preparada la solución, hay tres tipos de simulaciones. Las tres son ejecutadas de la misma manera: primero se ejecuta el comando "*grompp*" y a continuación el comando "*mdrun*".

El comando "*mdrun*" es el motor de GROMACS, responsable del verdadero desplazamiento de los átomos conforme a las leyes de la física (parametrizadas en el archivo de dinámica molecular *.mdp*).

El comando "*grompp*" toma los archivos *.gro*, *.top* y *.mdp* para generar un archivo *.tpr* que es la entrada del comando "*mdrun*".

Como salida del comando "*mdrun*" obtenemos un archivo binario *.trr* de gran tamaño, que contiene el estado del sistema a intervalos de tiempo regulares y también un archivo *.gro* que contiene el estado final de la solución. Este archivo *.gro* puede utilizarse para la simulación.

Hay tres tipos de simulaciones, aunque las dos primeras son opcionales, y realmente su función es la preparación de la solución para la tercera simulación, la completa.

La primera simulación tiene como objetivo la minimización de energías (*energy minimization* o EM) en la solución. Lo único que hace es "empujar" los átomos de las moléculas disueltas hasta que las longitudes de sus enlaces y sus ángulos se encuentren en su configuración de mínima energía potencial (ignorando por completo los átomos de otros elementos de la solución). Hacer esto es una buena manera de comprobar que los enlaces son estables (la molécula podría romperse si no lo fueran), y acelera las simulaciones posteriores.

En algunas soluciones son necesarios varios pasos de minimización de energía (EM), con lo cual será necesario ejecutar varias veces el ciclo *grompp->mdrun*.

La segunda simulación tiene por objetivo la fijación de la posición del soluto dentro de la disolución. Esto fija las moléculas de soluto en una posición, y permite que el disolvente y los iones se "relajen" a su alrededor en posiciones de mínima energía potencial. Con ello se consigue que el disolvente llene la caja de manera uniforme y no se formen vacíos al inicio de la simulación completa. Diseño de un simulador de dinámica molecular basado en CORBA. GROMACS.

La tercera simulación es la completa (*full molecular dynamics* o DM) y en ella se realizan los cálculos completos de la dinámica molecular de la solución en un periodo largo de tiempo

(las simulaciones anteriores necesitan mucho menos tiempo en comparación). Esta simulación es la que más tiempo de proceso necesita y es la más indicada para la realización en paralelo.

Como resultado de la simulación se obtienen archivos *.xtc*, *.trr* y *.edr* que son utilizados para el análisis de la simulación con utilidades como "*ngmx*" o "*xmgrace*", entre otras.

Un esquema completo de todas estas etapas y sus opciones se encuentra una vez instalado GROMACS, y generados los archivos necesarios para la simulación. La ejecución en secuencial de simulaciones se realiza con el comando *"mdrun*".

Este comando recibe múltiples parámetros, cuya descripción está documentada en la web de proyecto GROMACS.

Para las simulaciones realizadas con este proyecto, una vez realizados los pasos indicados anteriormente, la llamada al comando sería de la siguiente forma:

mdrun -v -s pr -e pr -o pr -c after_pr -g prlog

Esta llamada se realizaría sobre el directorio donde se ubiquen los archivos de la simulación

Función de distribución radial (RDF)

La función de distribución radial (RDF por sus siglas en inglés) $g(r)$ en un sistema de partículas (átomos, moléculas, etc.) está directamente relacionado con la probabilidad de encontrar otras moléculas, en función de la distancia medida desde una partícula de referencia. Considera una capa esférica de espesor δr a una distancia r de un átomo elegido. La $g(r)$ da la probabilidad de encontrar una partícula a una distancia "r" de otra partícula. Si contamos la aparición de dos moléculas a una separación r, de r = 0 a "R = ∞", se puede obtener la función de distribución radial g(r). Esta función es una herramienta útil para describir la estructura de un sistema, particularmente de líquidos. En un sólido, la distribución radial tiene un número infinito de picos agudos cuyas separaciones y alturas son características de la estructura reticular.

Disponibilidad del programa GROMACS

El programa más evolucionado para realizar una dinámica molecular clásica es GROMACS (Groningen Machine for Chemical Simulations), el cual está disponible en forma gratuita y fue diseñado para trabajar en sistemas operativos Unix y Linux. GROMACS; es una

aplicación de alto rendimiento que fue implementada y desarrollado por el departamento de Biophysical Chemistry de la Universidad de Groningen, y posteriormente ha sido reimplementada como código abierto por un amplio número de desarrolladores.

Paquetes de *software* para cálculos de estructura electrónica

Cálculos *ab initio* con la ecuación de Schrödinger

El término *ab initio* proviene del latín y significa 'desde el principio'. Se da este nombre a las propuestas computacionales derivadas directamente de principios teóricos (tales como la ecuación de Schrödinger), sin incluir información experimental. Las aproximaciones usuales incluyen el uso de una forma apropiada de una función cuando se aplica el método variacional, u obtener una solución aproximada a una ecuación diferencial (que implícitamente tiene errores de corte o de redondeo en el manejo de números).

El tipo más común de cálculo *ab initio* se conoce como cálculo Hartree-Fock (HF), en el cual la aproximación principal es la de campo central. Este es un cálculo variacional, lo cual significa que las energías aproximadas calculadas son iguales o más grandes a la energía exacta. Debido a la aproximación de campo central, las energías de los cálculos HF son siempre más grandes que la energía exacta y tienden al valor límite de Hartree-Fock. El método de Hartree-Fock depende de promedios, es decir, no considera las interacciones electrostáticas instantáneas entre los electrones y no toma en cuenta los efectos mecánico-cuánticos sobre las distribuciones de electrones, porque el efecto de N-1 electrones sobre un electrón de interés se trata en forma promedio.

Una medida de la aproximación de las soluciones que proporciona el método HF a la descripción de un sistema molecular es la energía de correlación de Löwdin (1958), que se define como: la diferencia de energía de correlación para un cierto estado con respecto a un Hamiltoniano es la diferencia entre el valor propio del Hamiltoniano exacto y su valor esperado en la aproximación de Hartree-Fock, para el estado en consideración.

Es necesario resaltar que el método de HF proporciona un número grande de resultados valiosos acerca de la estructura electrónica de las moléculas. Sin embargo, es una aproximación y no incluye correlación electrónica; proporciona un conjunto finito de espín-orbitales cuando se lleva a cabo una expansión en un conjunto base finito. En general, una base con M miembros resulta en 2M diferentes espín-orbitales. Se ordenan los espín-orbitales

ocupados de acuerdo con su energía, por los N electrones, y su función de onda de Hartree-Fock correspondiente, y los 2M-N orbitales virtuales permanecen ahí de acuerdo con el determinante de Slater.

La mayoría de los cálculos *ab initio* comienzan con un cálculo HF, seguido por cálculos para la repulsión electrón-electrón explícita, denominados correlación electrónica, entre los que destacan:

- La teoría perturbacional Möller-Plesset (MP)
- El método generalizado enlace valencia (GVB)
- Método multiconfiguracional de campo autoconsistente (MCSCF)
- Interacción de configuraciones (CI)
- La teoría de clúster acoplado (CC)

El aspecto favorable de los métodos *ab initio* es que eventualmente convergen en la solución exacta, una vez que todas las aproximaciones se han hecho suficientemente pequeñas en magnitud (con tolerancia < 10^{-6}). Sin embargo, a veces el cálculo más pequeño produce el mejor resultado para una cierta propiedad.

El aspecto desfavorable de los métodos *ab initio* que resuelven la ecuación de Schrödinger, es que son caros computacionalmente. A menudo toman bastante tiempo de CPU, memoria y espacio de disco. El método HF escala 4N, donde N es el número de funciones base, así que un cálculo del doble de tamaño toma 16 veces más tiempo para completarse. Los cálculos correlacionados se escalan aún peor que esto. En la práctica, soluciones extremadamente certeras se obtienen solo cuando la molécula contiene media docena de electrones o menos. A continuación, se menciona una lista de algunos de los paquetes de cómputo con los que se pueden realizar interacciones moleculares cuánticas.

Lista de programas computacionales

HONDO

PSHONDO-IJKL-GMCP-FOK-CIPSI

GAMESS

Gaussian

NWChem

QUANTUM EXPRESSO

VASP

BIOVIA Materials Studio

Rol de la computación en la química

Los grandes avances tecnológicos se deben a la gran acumulación de conocimientos que se emplean en diferentes áreas: automotriz, aviación, microelectrónica, industrias químicas e industrias biológicas, medicina, etc. Esto hace que la computación sea una herramienta importante. Los avances científicos teóricos y experimentales de los últimos años han incidido en computación, la cual se ha convertido en una herramienta muy potente, tal que actualmente permite el estudio de problemas que son cada vez más cercanos a la realidad y de un tamaño de partículas cercano a los ensayados en laboratorios de síntesis.

Son conocidas las leyes que gobiernan el comportamiento de iones y electrones en un sólido, un fluido o en una molécula, y son la base para resolver la ecuación de Schördinger. El desafío es hallar esta solución no solo para el caso del átomo de hidrógeno unielectrónico, sino para el problema multielectrónico de iones y átomos casi infinitos que forman un sólido. Un problema fundamental en esta dirección es el uso de simetrías, ya que son las que permiten obtener soluciones a la ecuación de Schrödinger y conducen a sistemas extremadamente grandes cuando se requiere resolver sistemas de un gran número de átomos. Una alternativa a la ecuación de Schrödinger es el uso de las ecuaciones de Kohn Sham, la cual se utiliza mucho para resolver sistemas moleculares con un mayor número de átomos, y que tiene sus propias limitaciones.

Si se presenta determinada complejidad del problema a resolver, combinada con la necesidad de lograr alta precisión, se tienen dos consecuencias importantes:

(a) los cálculos siempre tienen carácter de experimento computacional, y

(b) la estabilidad, precisión y eficiencia del algoritmo computacional son de suma importancia.

Dadas estas premisas, el rol de la computación en la química tiene un papel preponderante por las siguientes razones:

1. La computación permite aplicar la teoría a un problema complejo, es decir, un experimento de laboratorio nos dice qué ocurre, la teoría combinada con la informática posibilita encontrar por qué ocurre y podrá predecir qué llegará a pasar bajo circunstancias similares.
2. Complementar el experimento realizado en el laboratorio. Sucede cuando los experimentos son fácilmente accesibles por computadora, pero muchas veces no pueden ser abordados desde las mediciones experimentales de laboratorio.

   Se han llevado a cabo soluciones a la ecuación de Schrödinger para sistemas de pocos átomos interactuantes mediante los siguientes conjuntos de programas acoplados.

Programas acoplados secuencialmente

El conjunto de programas acoplados PSHONDO-IJKL-FOCK-CIPSI o PSHONDO-IJKL-GMCP-FOCK-CIPSI (desarrollados en lenguaje Fortran) ha permitido realizar una amplia variedad de estudios de reacciones de átomos metálicos con gases. Se ejecutan en sistema operativo Linux, no requieren requisitos previos para su ejecución. A continuación mostramos una descripción breve de cada uno de ellos.

i) Programa PSHONDO

Este es el programa computacional que realiza cálculos de sistemas moleculares. Se fundamenta en el método de campo autoconsistente de Hartree-Fock-Roothaan, además puede utilizar la aproximación matemática de pseudopotenciales dentro del esquema no empírico, desarrollado por Barthelat. Realiza cálculos de capas abiertas y/o de capas cerradas, es decir, hace el cálculo de sistemas atómicos y/o moleculares cuyos orbitales estén parcial o totalmente ocupados.

Para sistemas de capa cerrada, la función de onda se puede aproximar por un determinante de Slater y se realizan los cálculos usando el método HF-SCF. El cálculo desarrollado por este programa requiere inicialmente de un vector de entrada (función de onda de prueba) a partir del cual comienza el cálculo Hartree-Fock-Roothan, y al final se obtiene un vector o función de onda mejorada.

PSHONDO proporciona como resultado la energía electrónica total del sistema, las energías de los orbitales moleculares obtenidos, un análisis de población electrónica de cada

uno de los orbitales moleculares obtenidos y de cada centro, la carga total de cada átomo, el momento dipolar creado cuando existen desplazamientos netos de carga y los resultados de cada ciclo de los cálculos HF hasta lograr la convergencia.

Está diseñado para procesar información de hasta 30 átomos con 70 capas para 240 funciones gaussianas primitivas y 120 funciones base contraídas además de que usa los 5 orbitales tipo d (XX-YY,ZZ,XY,XZ,YZ) . Este programa requiere de las funciones base que describen a los electrones de valencia de cada uno de los átomos del sistema molecular estudiado, así como los parámetros correspondientes a los pseudopotenciales de los átomos que lo requieran en caso de que se vayan a utilizar. Para el conjunto de funciones base existe la opción de realizar contracciones entre las que son del mismo tipo. Dichas contracciones consisten en reemplazar las funciones base individuales por combinaciones lineales apropiadas. Esta información explícitamente se debe proporcionar a través de un archivo plano de entrada que será descrito más adelante.

ii) Programa IJKL

El programa IJKL emplea mucha de la información almacenada durante el proceso del programa PSHONDO, como las integrales atómicas mono y bielectrónicas que las transforma en integrales expresadas en términos de una base de orbitales moleculares. Además, ordena los orbitales moleculares de acuerdo con su simetría, la cual está asociada a las diferentes representaciones irreducibles del grupo puntual al que pertenece la molécula. Finalmente, en el interior de cada simetría, los orbitales moleculares se numeran según el orden energético creciente.

Este programa tiene valores definidos que no son parametrizables, entre los que destacan el uso de 20 representaciones irreducibles del grupo puntual como máximo, y un máximo de 120 orbitales atómicos de la base. Está condicionado a que se ejecute el programa completo y que la ejecución previa sea la del programa PSHONDO. También realiza la estimación del número de orbitales ocupados a partir de la fórmula siguiente:

$$\text{Número de Orbitales Ocupados} = (NE + 1) / 2$$

donde: NE es el número de electrones que se obtuvieron en el programa PSHONDO

Además, recibe como entrada 3 archivos que genera el programa PSHONDO

1.- El que contienen las integrales atómicas generadas.

2.- Los orbitales moleculares calculados por SCF.

3.- Los datos obtenidos, como el total de capas, el total de funciones base, el número de electrones, carga de la molécula, estado de la multiplicidad, número de orbitales ocupados, número de átomos, número de ejes de simetría, y grupo de simetría que serán reutilizados al inició de la ejecución del programa IJKL.

iii) Programa FOCK

Este programa junto con el programa IJKL sirven de encadenamiento entre el cálculo HF de PSHONDO y el proceso de CI que se lleva a cabo con el programa CIPSI.

El programa FOCK cierra las capas y calcula las energías orbitales, las cuales posteriormente se utilizan en la parte perturbativa del programa CIPSI. Además, diagonaliza la matriz del operador de Fock para el caso de capa cerrada, es decir, cuando los orbitales están doblemente ocupados. Entonces como resultado se tiene una matriz diagonalizada de $m \times m$ donde m es el número total de orbitales moleculares doblemente ocupados del sistema. Mediante este cálculo se define el determinante de referencia que sirve de partida para la inclusión de los efectos de correlación a nivel variacional y perturbativo, a través del uso del programa CIPSI (Bravo, 1986).

iv) Programa CIPSI

Este método efectúa correcciones perturbativas a segundo orden en forma iterativa, sobre funciones multiconfiguracionales obtenidas a través del procedimiento de interacción y configuraciones (CI).

El programa CIPSI tiene como función principal recuperar en buena medida la energía de correlación electrónica que no se contempla en los cálculos HF, esta energía se calcula a través de interacción de configuraciones con cálculos multireferenciales (MRCI) a niveles variacionales y perturbativos de segundo orden en Moller-Plesset; para que se lleven a cabo estos cálculos, la función de onda se expresa como una combinación lineal de determinantes de Slater que se representan en base a las excitaciones electrónicas a partir de un estado de referencia de capa cerrada (obtenido en el programa Fock) (Pacheco, 2009).

Por lo tanto, una vez que se tienen los orbitales moleculares etiquetados con su simetría correspondiente (generados en IJKL), se puede estudiar cualquier estado excitado

posible cuya simetría sea la de interés. Este programa proporciona los estados excitados que correlacionan con el estado estudiado. De estos el programa selecciona los determinantes de mayor contribución, de modo que éstos estados forman un espacio configuracional, con el cual el programa realiza un cálculo variacional de la energía a través de la diagonalización de la matriz asociada al hamiltoniano.

Los determinantes excitados con respecto a esta referencia se generan al sustituir uno o más espín-orbitales ocupados por un número equivalente de espín-orbitales virtuales compatibles con la simetría del espacio y espín deseado (Bravo, 1986).

Este sistema de programas acoplados para llevar a cabo cálculos *ab initio* está diseñado para una dimensión de hasta 200 configuraciones, un total de 70 orbitales, 20 estados perturbativos y 600 determinantes. De este sistema de programas se obtienen los valores del determinante, la norma, diagonalización, perturbación y energía total.

El programa CIPSI consiste en una selección iterativa de una primera clase de determinantes {S}, cuyo peso en la función de onda exacta Ψ_M es mayor que un cierto umbral η

$$D_K\ pertenece\ a\ \{S\} \quad si \quad |\langle \Psi_M | D_K \rangle| > \eta$$

la cual es una aproximación ampliamente usada para estudiar estados base y estados excitados. Esta primera clase de determinantes es tratada variacionalmente

$$P_S = \sum_{K\{S\}} |D_K\rangle\langle D_K|$$

P_S es la proyección sobre el subespacio S. CIPSI es un programa de Interacción de Configuraciones en el formalismo hueco-partícula a partir de un estado de referencia a capas completas ϕ_0 (determinante de referencia). Si *noc* es el número de orbitales ocupados entonces ϕ_0 se construye con $2 \times noc$ espinorbitales:

$$\phi_0 = |1\bar{1} \cdots i\bar{\imath} \cdots noc\overline{noc}|$$

los orbitales no son forzosamente aquellos que diagonalizan el operados de Fock correspondiente a capas completas

$$F = h + \sum_{j\ \epsilon\ oce} 2\left(J_j - K_j/2\right)$$

más bien la energía de referencia $\langle \phi_0 | H | \phi_0 \rangle$ calculada en el programa IJKL corresponde a estos operadores. Los determinantes D_K excitados con relación a esta referencia pueden

generarse sustituyendo uno o más espinorbitales ocupados (huecos) por otros tantos de espinorbitales virtuales (partículas) compatibles con la simetría del espacio y del spín deseado.

Primero se realiza un CI variacional dentro de un subespacio restringido de determinantes de simetría adecuada (espacio S) de dimensión NCF. Así obtenemos bajo el estado M considerado una función de onda multiconfiguracional Ψ_M^0 y una energía variacional ϵ_M tales que

$$P_S H P_S |\Psi_M^0\rangle = E_M^{(0)} |\Psi_M^0\rangle$$

$|\Psi_M^0\rangle$ se desarrolla sobre los determinantes de S, entonces la función de onda multiconfiguracional resultante es

$$|\Psi_M^0\rangle = \sum_{K \epsilon S} C_{MK}^0 \, |D_K\rangle$$

Se puede así definir un hamiltoniano de orden cero:

$$H_0 = \sum_{M=1}^{NCF} E_M^0 \, |\Psi_M^0\rangle\langle\Psi_M^0| + |D_I\rangle\langle D_I|$$

D_I es un determinante exterior a S. Análogamente se pueden calcular las contribuciones de orden 1 a la función de onda:

$$|\Psi_M^1\rangle = |\Psi_M^0\rangle + \sum_{I \notin \{S\}} \frac{\langle \Psi_M^0 | H | D_I \rangle}{E_M^0 - E_I^0} |D_I\rangle$$

se le llama C_{MI} a la contribución

$$C_{MI} = \frac{\langle \Psi_M^0 | H | D_I \rangle}{E_M^0 - E_I^0}$$

Los resultados dependen obviamente del tipo de partición utilizado. Es así perturbada a segundo orden por los demás determinantes que no pertenecen a $\{S\}$

$$E_M^{(2)} = \sum_{I \notin \{S\}} \frac{\langle \Psi_M^0 | H | D_I \rangle^2}{E_M^0 - E_I^0}.$$

En la práctica la definición de $\{S\}$ es iterativa: comenzando en una propuesta inicial de $\{S\}$ uno elige los determinantes más importantes de la corrección a primer-orden de la función de onda. Si el coeficiente a primer-orden de D_I es mayor que η^k en la k-ésima iteración

$$\left|\frac{\langle \Psi_M^0|H|D_I\rangle}{\Delta E}\right| > \eta^k$$

D_I es agregado a $\{S\}$ y el proceso se puede repetir decreciendo η^k a $\eta^{k+1} < \eta^k$. En tal caso no hay una selección *a priori* del espacio de multireferencia $\{S\}$, cuya calidad se incrementa progresivamente.

La secuencia de programas 'psn → ijkl → fok → gmcp → cipsi' para un conjunto de bases se comienza a correr con la instrucción

%psn <file.dat >file.psn &

Con el fin de encontrar la energía de CIPSI

$$E_{\text{CIPSI}} = E_{\text{SCF}} + E_{\text{DIAG}} + E_{\text{PERT}}$$

a partir de la energía de Hartree-Fock

$$E_{HF} \rightarrow \begin{Bmatrix} P_S H P_S|\Psi_m^0\rangle = \epsilon_m^{(0)}|\Psi_m^0\rangle \\ \Psi_m^0 = \sum_K C_{mK}\,\Phi_K \end{Bmatrix} \rightarrow \epsilon_m^{(2)} - \sum_{I\notin\{S\}} \frac{\langle \Psi_m^0|H|\Phi_I\rangle^2}{E_m^0 - E_I^0}.$$

Procedimiento de cálculos ab initio PSHONDO-IJKL-GMCP-FOCK-CIPSI

La ejecución del conjunto de programas acoplados inicia con el programa *PSHONDO-IJKL-GMCP-FOCK-CIPSI* a partir de un archivo de entrada que se caracteriza por secciones definidas a través de etiquetas con su respectivo conjunto de variables, como se describe a continuación:

&HONDO: Variables generales del programa PSHONDO.

IUNT= Tipo de unidad Si el valor es <=0 indica unidades atómicas; >0 para angstrom.

POP= *True* si el análisis de la población electrónica es requerido.
False de lo contrario.

GROUP= Símbolo Schoenflies del grupo de simetría puntual. Los posibles símbolos son C1 (Valor por *Default*), CS, CI, CN, S2N, CNH, CNV, DN, DNH, DND, CINFV, DINFH, T, TH, TD, O, OH, I, IH.

NAXIS= Orden de los principales ejes de simetría.

NPRINT= 0: Impresión normal de los resultados de todos los cálculos.
1: Impresión únicamente de todas las integrales.
2: Impresión extensa solamente del conjunto de base.

3: Impresión extensa del conjunto de base y datos de simetría.

4: Impresión simple de 72 caracteres por línea.

5: Impresión solamente de datos SCF después de cada ciclo.

6: Impresión de análisis de población de cada orbital molecular.

Valor *default* 4.

TIMLIM= Límite de tiempo interno que para el programa. El tiempo de CPU es obtenido por el reloj del sistema y después es comparado con este valor. Valor *default* 1750 segundos.

WIJKL= *True* si se ejecutará el programa IJKL al finalizar el PSHONDO. *False* de lo contrario.

MULTIP= Multiplicidad. 1 para capa cerrada.

&END

&CENT: Variables con datos de las bases atómicas y geometría de la molécula

ATOM= Nombre atómico.

IATN= Número atómico.

ZNUC= Carga nuclear.

NTYPG= Vector de la base Gaussiana contraída. Se define por el tipo de la capa como sigue:

0: para capa s

1: para capa p o l

2: para capa d

NCONG= Vector de primitivas contraídas en cada capa.

EXG= Vector de exponentes gaussianos.

CSG= Vector de coeficientes de contracción para las funciones s.

CPG= Vector de coeficientes de contracción para las funciones p.

CDG= Vector de coeficientes de contracción para las funciones d.

X, Y, Z= Coordenadas para cada eje respectivamente.

&END

&PSDINP: Variables del pseudopotencial atómico. Solo si la variable PSEUDO= *True* (esta variable está en código duro en el programa PSHONDO y que por *default* tiene ese valor).

NUMAT= Número atómico del átomo que genera el pseudopotencial.

NBTYP= Vector con los números de términos en cada simetría.

APOT= Vector de exponentes de la parte exponencial.

CPOT= Vector de los coeficientes.

NPOT= Vector de las potencias de R.

&END

&SCFINP: Lista de variables para la convergencia de SCF.

NPUNCH= Valor default 0. Indica orbitales de energía y orbitales ocupados.

BNAME= Si el valor es PSEUDO indica que se realizan cálculos de tipo SCF Nesbet.

ANAME= VECTOR que indica que se dan vectores de prueba.

QN = Número de orbitales ocupados (Si BNAME=PSEUDO). Valor default 2.

ALPHA= Coeficiente utilizado para formar la nueva matriz de densidad. Valor default 1.0.

NCONV= Criterio de convergencia. La convergencia se cumple cuando la variación de la energía entre dos ciclos SCF consecutivos es menor que el valor de $10.0^{(NCONV)}$ en valor absoluto. Valor *default* 6.

MAXIT= Número máximo de iteraciones. Valor *default* 30.

&END

&VECINP: Es requerido si los vectores de prueba se proporcionaron directamente dado que la variable ANAME=VECTOR

IVEC= Número positivo entero que especifica el archivo donde se leen los Coeficientes.

&END

&IJKLIN: Variables para la preparación de los datos que utilizará el programa ijkl.

YORBF= F Indica que se ejecutó anteriormente el programa PSHONDO. El valor T es para indicar que se ejecutó el programa PSHF que no está considerado en el alcance de este trabajo.

YRECSP= T Para indicar que se obtengan las degeneraciones de los orbitales, así como su simetría.

NTRANS= 0 Indica la transformación de la simetría.

YPRT= F Cuando no se requiere obtener el resumen de los cálculos realizados en el programa PSHONDO.

TDEG= Valor usado para la determinación de las degeneraciones.

TREC= Valor usado para verificar la ortogonalización.

YWVECT= T si se obtiene del archivo la matriz completa de los orbitales moleculares por representaciones irreducibles.

YF11= T Si existe el archivo que contiene los orbitales moleculares del cálculo SCF del programa PSHONDO.

&END

&ICINP: Variables para la preparación de los datos que utilizará el programa CIPSI.

TITLE= Nombre de la molécula y o átomos involucrados en el cálculo.

METAT= Número de estados que están perturbados. Valor *Default* 1.

NUMERO= Nombre de estados que están perturbados.

YION= F (valor por default) Cuando la excitación en l no es infinita.

ISZ= 0 (valor *default*) Valor del estado S_z

YPERTU= T Realiza cálculos de perturbación Moller-Plesset baricéntrica.

YENVP= F (valor *default*) No calcula la perturbación Moller Plesset.

YENB= F (valor *default*) No calcula la perturbación Epstein-Nesbet.

TEST= Umbral absoluto de la selección de determinantes por cada estado en la perturbación Epstein-Nesbet. Valor *Default* 0.5.

YKEEP= F Para no guardar en archivo los resultados variacionales de energía.

YMOYEN= F Para no escribir en archivo el histograma de la perturbación.

TAU= Umbral de la selección en la escritura en archivo de determinantes para el histograma. (Solo si la variable YMOYEN=T).

NCLASS= Número de límites del histograma, los valores se definen en la variable AVAL. Valor *default* 10. (Solo si la variable YMOYEN=T)

AVAL= Valores de los límites del histograma. (Solo si YMOYEN=T)

&END

Ejemplo: Entrada para calcular estados de energía de níquel

```
&HONDO
     IUNT=0,
     POP=.TRUE.,
     GROUP='CNV', NAXIS=2,
     NPRINT=0,
     TIMLIM=80000.,
     WIJKL=.TRUE. ,
     MULTIP=3,   &END
 &CENT ATOM='NI', IATN=28, ZNUC=10.0,
     NTYPG=3*0,3*1,3*2,
     NCONG=1,1,1, 1,1,1, 3,2,1,
     EXG=0.689923,0.156842,0.052230,
        1.199960,0.218215,0.068306,
        32.880202,9.028273,2.944905,1.039471,0.378960,0.139114,
     CSG=-0.254483,0.533751,0.637711,
     CPG=-0.048703,0.535795,0.573695,
     CDG=0.053138,0.240617,0.442008,0.382806,0.200701,0.045905,
     X=0.0, Y=0.0, Z=0.0,
 &END
 &CENT       &END
 &PSDINP NUMAT=28, NBTYP=0,2,1,4,
       APOT=2*1.021538,1.400000,4*2.738530,
       CPOT=10.550132,16.418257,
           -0.581050,
           -1.012088,-6.353834,6.655059,-11.008880,
```

```
      NPOT=-1,-2, -2, 0,-1,-2,0,
 &END
 &SCFINP NPUNCH=-1, BNAME='PSEUDO', ANAME='VECTOR',
      QN=6*1.333,1,1, ALPHA=0.8, NCONV=7, MAXIT=30, &END
 &VECINP ivec=12,
 &END
 &IJKLIN yorbf=f, yrecsp=t, ntrans=0,
       yprt=f, tdeg=1.0d-7, trec=1.0d-7, ywvect=t,
       yf11=t,
 &END
 &MCINP nconv=4, irest=0, mspin=3, iskip=0,
   idiagn=2, idp=1, iprinm=1, tdeg=1.0d-7, niter=60,
 &END
 &conf idoc=1,1,3,5, isoc=2,2,6,6,   &end
 &conf    &end
 &ICINP TITLE='ni',
  TOCOM=2,1,3*2,1,
  METAT=1, NUMERO=1,
  yion=f, isz=0,
  yPERTU=t, yENVP=f, yENB=f, TEST=0.01, yKEEP=f,
  yMOYEN=f, TAU=0.001, nclass=9,
  aval=0, 1.d-4, 1.d-3, 3.d-3, 5.d-3, 7.d-3, 9.d-3, 1.1e-2, 1.,
 &END
M 2 6
 &MOYEN TAU=1.0D-3, MAXDET=30000, &end
```

Ejemplo: en el caso del átomo de oro, los valores numéricos de las bases se consiguen de la siguiente manera

$$^{79}\text{Au}\ [\text{Xe}]4f^{14}5d^{10}6s$$

FCC a = 4.078 Å

Átomo	Configuración	Valencia	Esquema de contracción
Au	[Xe]$4f^{14}$	$5d^{10}6s^{1}$	(3s3p4d)/[111/21/211]
C	[He]	$2s^{2}2p^{2}$	(4s4p1d)/[211/211/1]
N	[He]	$2s^{2}2p^{3}$	(4s4p1d)/[211/211/1]
O	[He]	$2s^{2}2p^{4}$	(4s4p1d)/[211/211/1]

^{6}C [He]$2s^{2}2p^{2}$ Dunning d(0.75)

^{7}N [He]$2s^{2}2p^{3}$ Dunning d(0.80)

^{8}O [He]$2s^{2}2p^{4}$ Dunning d(0.85)

Bases de R.B. Ross, W.C. Ermler, P.A. Christiansen [75]

	H – He	**Li – Ne**	**Na - Ar**
cc – pVDZ	[2s1p] → 5 func	[3s2p1d] → 14 func	[4s3p1d] → 18 func
cc – pVTZ	[3s3p1d] → 14 func	[4s3p2d1f] → 30 func	[5s4p2d1f] → 34 func
cc – pVQZ	[4s3p2d1f] → 30 func	[5s4p3d2f1g] → 55 func	[6s5p4d3f2g1h] → 59 func

Bases	**Funciones primitivas**	**Funciones contraídas**
cc – pVDZ	9s,4p,1d/4s,1p	3s,2p1d/2s,1p
cc – pVTZ	10s,5p,2d,1f/5s,2p,1d	4s,3p,2d,1f/3s,2p,1d
cc – pVQZ	12s,6p,3d,2f,1g/6s,3p,2d,1f	5s,4p,3d,2f,1g/4s,3p,2d,1f
cc – pV5Z	14s,9p,4d,3f,2g,1h/8s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h/5s,4p,3d,2f,1g
cc – pV6Z	16s,10p,5d,4f,3g,2h,1i/10s,5p,4d,3f,2g,1h	7s,6p,5d,4f,3g,2h,1i/6s,5p,4d,3f,2g,1h

```
$HONDO
    IUNT=1,
    GROUP='CNV', NAXIS=2,
    NPRINT=0,
    WIJKL=.TRUE.,
    TIMLIM=80000.,
    MULTIP=2,
$END
 $CENT ATOM='Au', IATN=79, ZNUC=11.0,
     NTYPG=3*0,2*1,3*2,
     NCONG=1,1,1,2,1,2,1,1,
     EXG=0.4409000,0.2626000,0.0617000,
        0.6642000,0.1073000,0.0319000,
       1.3517000,0.5149000,0.1994000,0.0776000,
     CSG=-0.7478000, 0.8176000, 0.8113000,
     CPG=-0.0459000, 0.3778000, 0.7233000,
     CDG=0.4787000, 0.4582000, 0.2203000, 0.0412000,
    X=0.0, Y=0.0, Z=0.0,
$END
```

En este caso, los datos de los coeficientes de contracción de las funciones s, p, d y el vector de exponentes se obtienen de la referencia (Ross et al., 1990).

5d,6s,6p Valence Space [75]			
C (n=6)	Exp.	C (n=5)	C (n=4)
(-0.2909)			(-0.2927
1.2599	0.4409		-0.7478
-1.5250	0.2626		0.8176
-0.2533	0.0617		0.8113
0.6171			
0.5872			
			(-0.1274)
	0.6642		-0.0459
	0.1073		0.3778
	0.0318		0.7233
		(-0.4516)	
	1.3517	0.4787	
	0.5149	0.4582	
	0.1994	0.2203	
	0.0776	0.0412	

Pseudopotenciales

Los pseudopotenciales atómicos se utilizan para simular los efectos debidos a los electrones del carozo atómico, y de esta manera simplificar el problema electrónico molecular. Se requiere de un cálculo con pseudopotenciales que reproduzca correctamente el espectro monoelectrónico correspondiente a los orbitales de valencia de átomos diferentes, los cuales son responsables de la formación o no de enlaces químicos. Esto se expresa a través de la condición:

$$F_i^{ps}\phi_{\mathrm{V}} = \varepsilon_{\mathrm{V}}\phi_{\mathrm{V}}$$

En donde F_i^{ps} es el operador monoelectrónico de Fock, ε_{V} es la energía orbital de un cálculo con todos los electrones relacionada con el orbital verdadero de valencia ψ_{V}, y ϕ_{V} es el

pseudo-orbital de valencia. Con el fin de reproducir las propiedades moleculares, el pseudo-orbital ϕ_V debe ser prácticamente idéntico al orbital verdadero ψ_V en la región de valencia. Durand y Barthelat (1975) la determinaron extendiéndose más allá del radio del carozo R_C que corresponde al punto de intersección entre el orbital del carozo de la capa inmediatamente abajo y el orbital de valencia verdadero ψ_V.

$$W = \sum_{\ell=0}^{\ell_{max}} W_\ell \hat{P}_\ell \, (\theta_i, \phi_i)$$

En donde $\hat{P}_\ell$ es el proyector sobre los armónicos esféricos de orden ℓ, y la parte radial está dada por

$$W_\ell = e^{-\alpha r^2} \sum_{i=1}^{n} C_i r^{n_i} \qquad \ell = 0,1,2$$

Los nombres de las variables corresponden a entradas en el programa PSHONDO tales que

NBTYP = n,

APOT = α,

CPOT = C_i

NPOT = n_i

Una base gaussiana estándar en coordenadas cartesianas se optimiza sobre el estado fundamental del átomo con el pseudopotencial precedente. Estas bases, generalmente formadas de cuatro primitivas, pueden contraerse.

```
$PSDINP NUMAT=6,8,79, NBTYP=0,7,0,0,0,7,0,0,0,14,14,14,14,
    APOT=3.81909999, 4.17320001, 6.27069998, 12.21120000,
        5.35280001, 18.06680012, 51.61590004,
        2.24790001, 2.40490001, 4.37400001, 2.18920001,
        10.02859998, 34.19799995, 100.00389957,
        1.22479999, 1.43359995, 1.89289999, 2.69630003, 3.98990011, 6.00549984,
14.58259964, 39.31069946,
        0.96569997, 2.29040003, 6.43839979, 16.36459923, 55.98929977,
171.10139465,
```

0.90640002, 1.03110003, 1.32869995, 1.90369999, 2.83640003, 4.58839989, 11.19810009, 32.29570007,

0.96569997, 2.29040003, 6.43839979, 16.36459923, 55.98929977, 171.10139465,

0.29310000, 1.03009999, 2.01859999, 3.43919992, 5.15789986, 2.51489997, 8.06239986, 41.02849960,

0.96569997, 2.29040003, 6.43839979, 16.36459923, 55.98929977, 171.10139465,

0.43349999, 0.50970000 0.57959998, 0.77109998, 2.59349990, 8.39070034, 6.92939997, 22.25889969,

0.96569997, 2.29040003, 6.43839979, 16.36459923, 55.98929977, 171.10139465,

CPOT=-47.09821510, 71.58925819, -4.67536402, 3.03797001,

-0.55931300, -4.07454997, -1.43484600,

11.21630394, -16.34447694, 1.04294400, 2.19389099,

-0.79842000, -5.76684701, -1.48645601,

-84.96929932, 280.17764282, -354.64355469, 334.33190918, -219.39260864, 170.85874939, 56.33872986, 6.40929985,

-0.98690200, -9.68630600, -32.57246399, -118.53087616, -286.25463867, -51.27423096,

-6.91190100, 72.12546539, -114.97533417, 123.61668396, -71.13571167, 79.01660156, 51.28700638, 5.65198278,

-0.98690200, -9.68630600, -32.57246399, -118.53087616, -286.25463867, -51.27423096,

-0.11365700, -2.40402508, 62.58111572, 255.44717407, -161.80647278, -175.77076721, 63.62079239, 6.75031996,

-0.98690200, -9.68630600, -32.57246399, -118.53087616, -286.25463867, -51.27423096,

-47.60779953, 176.31802368, -168.67689514, 54.49244690, 37.71539307, -24.13967514, 49.16802216, 4.75958586,

-0.98690200, -9.68630600, -32.57246399, -118.53087616, -286.25463867, -51.27423096,

NPOT=0,0,-1,-2,
0,0,-1,
0,0,-1,-2,
0,0,-1,
0,0,0,0,0,0,-1,-2,
0,0,0,0,0,-1,
0,0,0,0,0,0,-1,-2,
0,0,0,0,0,-1,
0,0,0,0,0,0,-1,-2,
0,0,0,0,0,-1,
0,0,0,0,0,0,-1,-2,
0,0,0,0,0,-1,
$END

Los valores de APOT resaltados en negritas fueron reportados por Ross et al. (1990). Dichos datos se mencionan enseguida

Atom	n_{1i}	Exp.	C_1i
Gold - 5d, 6s, 6p, Valence Space			
s - g	2	1.2248	-84.969303
	2	1.4336	280.177643
	2	1.8929	-354.643566
	2	2.6963	334.331909
	2	3.9899	-219.392605
	2	6.0055	170.858744
	1	14.5826	56.338731
	0	39.3107	6.409300
p - g	2	0.9064	-6.911901
	2	1.0311	72.125463

	2	1.3287	-114.975335
	2	1.9037	123.616682
	2	2.8364	-71.135709
	2	4.5884	79.016603
	1	11.1981	51.287008
	0	32.2957	5.651983
d - g	2	0.2931	-0.113657
	2	1.0301	-2.404025
	2	2.0186	62.581116
	2	3.4392	255.447176
	2	5.1579	-161.806480
	2	2.5149	-175.770760
	1	8.0624	63.620792
	0	41.0285	6.750320
f – g	2	0.4335	-47.607799
	2	0.5097	176.318029
	2	0.5796	-168.676897
	2	0.7711	54.492446
	2	2.5935	37.715395
	2	8.3907	-24.139676
	1	6.9294	49.168024
	0	22.2589	4.759586
g	2	0.9657	-0.986902
	2	2.2904	-9.686306
	2	6.4384	-32.572463
	2	16.3646	-118.530878
	2	55.9893	-286.254646
	1	171.1014	-51.274231

Disponibilidad de programas

El programa HONDO fue inicialmente elaborado en Francia para resolver problemas de física atómica y molecular. Uno de sus desarrolladores vino a una estancia en México, y junto con investigadores mexicanos se desarrolló una versión muy utilizada en cálculos pioneros de catálisis.

El programa HONDO se conoce en México como PSHONDO-IJKL-GMCP-FOK-CIPSI, y sirve para resolver problemas de pocos átomos. Una de sus debilidades es encontrar las bases apropiadas al sistema en estudio que permitan credibilidad en la solución, y la otra es que en México ha quedado rezagado su desarrollo en crecer a sistemas *full* CI. En la Universidad de Toulouse (en Francia), dicho programa ya ha sido extendido a cálculos *full* CI, y se desconoce su disponibilidad. Sin embargo, quien lo aprende a utilizar artesanal y eficientemente consigue tanto una base como una interacción de configuraciones apropiada. Esto se ha logrado para muchos sistemas que ya se han publicado para problemas *ab-initio* de sistemas pequeños de átomos y moléculas, y es muy útil para comparar su solución contra las de otros programas más evolucionados. Está disponible en plataforma Linux Red Hat para aquellos que lo conocen en algunas instituciones como: UAMA, IFUNAM y TecNM/ITT (en México).

GAMESS es otra versión del programa HONDO y es libre, está disponible en internet bajo su previa solicitud. No obstante, GAMESS se usa de una manera distinta a la forma en que se usa PSHONDO-IJKL-GMCP-FOK-CIPSI.

El programa Gaussian es uno de los más empleados, y fue libre durante mucho tiempo; sin embargo, ahora es un *software* de licencia. Es un *software* muy robusto que deja resolver problemas no solo de Física atómica y molecular, sino también de química cuántica dados los módulos de DFT que tiene integrados. También aparece en uno de los módulos de BIOVIA Materials Studio y, por ejemplo, dada su versatilidad para realizar computacionalmente análisis espectroscópicos, sirve para estudiar óxidos metálicos, o interacciones oxígeno-metal; sin embargo, no se puede ejecutar como cualquier otro programa de MS Modeling.

El *software* NWChem, que es un programa alternativo a Gaussian, nos ha permitido hacer cálculos de dinámica molecular Born-Oppenheimer, en los que los cálculos *single-*

point son superados imponiendo una velocidad a la partícula escogida para atacar la otra, aunque los resultados solo aparecen en función del tamaño de paso en femptosegundos. En el análisis de la dinámica cuántica, cada núcleo sigue una trayectoria newtoniana debido a las fuerzas derivadas del potencial del electrón más la fuerza electrostática ejercida por las otras partículas nucleares. El enfoque autoconsistente se considera convergido o terminado cuando se obtiene una tolerancia de 0.00001 tanto para la energía como para la densidad (Zaragoza et al., 2018).

QUANTUM EXPRESSO es un programa libre, que puede hacer lo que hace VASP aunque no tan rápidamente.

VASP es un programa de licencia que es muy útil para hacer cálculos entre carbono, gases y metales. Es muy apropiado para el tipo de cálculos que cualquier investigador dedicado a cálculos computacionales desearía tener.

BIOVIA Materials Studio es un programa (de licencia) interactivo muy versátil, con el cual se pueden realizar cálculos: semiempíricos, mesoscópicos, poliméricos, de teoría de funcionales de la densidad (DFT), de dinámica molecular (tanto clásica como cuántica), de dinámicas de partículas disipativas, y de estado sólido (CASTEP), al menos. Es muy apropiado para el tipo de cálculos que se desarrollan en el Instituto Tecnológico de Toluca, dado que es de fácil uso tanto para estudiantes como profesores. Su versatilidad incluye el poder hacer cálculos con computadoras de escasos recursos, y se consiguen cálculos más fácilmente conforme el poder de cómputo en hardware adquiere mayores dimensiones en recursos.

Anexos

comunicacion-cientifica.com

Anexo A. Relación fundamental de la termodinámica

Se dice que un sistema es *homogéneo* si las propiedades que lo describen permanecen invariables en toda su extensión espacial. En caso contrario el sistema es *inhomogéneo*. Se sabe que cantidades tales como el volumen V, la masa *m* y la fracción mole *n* de un sistema se llaman variables *extensivas*, ya que sus valores dependen de la cantidad de materia presente en el sistema (Prigogine y Defay, 1954), mientras que variables tales como la temperatura T, la presión p y la fracción mol x_i de la componente *i* son variables intensivas, ya que tienen valores definidos en cada punto del sistema. El valor de una variable intensiva puede ser el mismo en todo el sistema o cambiar de un punto a otro.

El principal objetivo de la termodinámica en el estudio de un sistema específico es determinar la ***r**elación **f**undamental de la **t**ermodinámica* (***R F T***). Para un sistema unicomponente se tiene que:

$$S = S[U, n, V], \tag{A1}$$

en donde S es la entropía, *n* es el número de moles o masa, U es la energía y V es el volumen. Si se conoce explícitamente la ***R F T,*** entonces es factible deducir el comportamiento del sistema en cualquier punto del espacio termodinámico (generado por la energía, la entropía, el volumen y el número de moles de la especie contenida en el sistema) (Callen, 1996). Es decir, toda la información relevante acerca del sistema estudiado está contenida en la ***R F T.*** La especificación de dicha ***R F T*** puede llevarse a cabo haciendo diversos niveles de descripción, dependiendo del interés y la complejidad del proceso que se desea estudiar.

A continuación, se verá de una forma muy simple cómo se puede deducir los potenciales termodinámicos a partir de la ***R F T***. Sobre la base de las dos primeras leyes de la termodinámica para procesos reversibles: 1ª $dU = dQ - PdV + \mu dn$, y 2ª. $dS = \frac{dQ}{T}$. Sustituyendo la primera en la segunda ley de la termodinámica se obtiene la expresión

$$dS = \frac{1}{T}\left(dU + PdV - \mu dn\right), \tag{A2}$$

y como la diferencial de la entropía está dado por

$$dS = \frac{\partial S}{\partial U}dU + \frac{\partial S}{\partial V}dV + \frac{\partial S}{\partial n}dn, \tag{A3}$$

entonces los parámetros extensivos están relacionados a los intensivos de la siguiente manera

$$\left(\frac{\partial S}{\partial U}\right)_{V,n} \equiv \frac{1}{T}, \tag{A4}$$

$$\left(\frac{\partial S}{\partial V}\right)_{U,n} \equiv \frac{P}{T}, \tag{A5}$$

$$\left(\frac{\partial S}{\partial n}\right)_{U,V} \equiv -\frac{\mu}{T}. \tag{A6}$$

De la ecuación (A2) se obtiene

$$dU = TdS - PdV + \mu dn, \tag{A7}$$

la cual es una forma alternativa de la ***R F T*** para la energía del sistema, que de manera general se puede expresar como

$$\mathrm{U} = \mathrm{U}[\mathrm{S}, \mathrm{V}, n]. \tag{A8}$$

Cuya diferencial está dada por

$$dU = \left(\frac{\partial U}{\partial S}\right)_{V,n} dS + \left(\frac{\partial U}{\partial V}\right)_{S,n} dV + \left(\frac{\partial U}{\partial n}\right)_{S,V} dn. \tag{A9}$$

Las derivadas parciales que aparecen en la ecuación anterior se conocen como los *parámetros intensivos.* Dichos parámetros quedan entonces definidos de la siguiente manera:

$$\left(\frac{\partial U}{\partial S}\right)_{V,n} \equiv T, \quad la \quad temperatura \tag{A10}$$

$$-\left(\frac{\partial U}{\partial V}\right)_{S,n} \equiv P, \quad la \quad presión \tag{A11}$$

$$\left(\frac{\partial U}{\partial n}\right)_{S,V} \equiv \mu, \quad el \quad potencial \quad químico. \tag{A12}$$

Este conjunto de variables intensivas se puede expresar en términos de los parámetros extensivos independientes S, V, *n*, por ejemplo, para un gas ideal monoatómico*:*

$$T = T_0\left\{\frac{V_0 n}{V n_0}\exp\left(\frac{1}{R}\left(\frac{S}{n}-\frac{S_0}{n_0}\right)\right)\right\}^{2/3}, \tag{A13}$$

en la cual se aprecia claramente que:

$$T = T(S,V,n), \tag{A14}$$

o bien también puede quedar como función de

$$P = P(S,V,n), \tag{A15}$$

o

$$\mu = \mu(S,V,n), \tag{A16}$$

las cuales se llaman ecuaciones de estado. Si las tres ecuaciones de estado son conocidas, pueden sustituirse en la ecuación (A7), recuperando así la ecuación fundamental al integrar dicha ecuación. Entonces, la totalidad de las tres ecuaciones de estado es equivalente a la ***R F T***, y contiene toda la información termodinámica acerca del sistema.

Si en la ecuación (A8) para la energía U se reemplaza la entropía S por la temperatura T bajo una transformada de Legendre, es decir de la ecuación (A7) se obtiene

$$dU = TdS + SdT - SdT - PdV + \mu dn, \tag{A17}$$

$$\therefore \quad d(U - TS) = -SdT - PdV + \mu dn \tag{A18}$$

de donde se tiene que la *energía libre de Helmholtz* está dada por

$$F = U - TS, \tag{A19}$$

donde

$$F = F(T,V,n), \tag{A20}$$

$$\left(\frac{\partial F}{\partial T}\right)_{V,n} \equiv -S, \quad la \quad entropía \tag{A21}$$

$$\left(\frac{\partial F}{\partial V}\right)_{T,n} \equiv -P, \quad la \quad presión \tag{A22}$$

$$\left(\frac{\partial F}{\partial n}\right)_{T,V} \equiv \mu, \quad el \quad potencial \quad químico. \tag{A23}$$

Análogamente, si en la ecuación (A8) de la energía U se reemplaza el volumen V por la presión P bajo la transformada de Legendre, se obtiene la *entalpía*

$$dH = TdS + VdP + \mu dn,$$

$$H = U + PV, \quad donde \quad H = H(S,P,n), \tag{A24}$$

$$\left(\frac{\partial H}{\partial S}\right)_{P,n} = T,$$

$$\left(\frac{\partial H}{\partial P}\right)_{S,n} = V,$$

$$\left(\frac{\partial H}{\partial n}\right)_{S,P} = \mu.$$

Una vez más, si en la ecuación (A8) de la energía U, y bajo la transformada de Legendre se reemplaza tanto la entropía S por la temperatura T, como el volumen V por la presión P, se obtiene la *energía libre de Gibbs*

$$dG = -SdT + VdP + \mu dn,$$

$$G = U - TS + PV, \quad donde \quad G = G(T,P,n). \tag{A25}$$

$$\left(\frac{\partial G}{\partial T}\right)_{V,n} \equiv -S, \quad la \quad entropía \tag{A26}$$

$$\left(\frac{\partial G}{\partial P}\right)_{T,n} \equiv V, \quad el \quad volumen \tag{A27}$$

$$\left(\frac{\partial G}{\partial n}\right)_{T,V} \equiv \mu, \quad el \quad potencial \quad químico. \tag{A28}$$

Nótese, además, que de la definición de la propiedad homogénea de primer orden se tiene que para cualquier λ

$$U(\lambda S, \lambda V, \lambda n) = \lambda U(S,V,n),$$

$$\left(\frac{\partial U(\lambda S,\lambda V,\lambda n)}{\partial \lambda S}\right)_{\lambda=1} \frac{\partial(\lambda S)}{\partial \lambda} + \left(\frac{\partial U(\lambda S,\lambda V,\lambda n)}{\partial \lambda S}\right)_{\lambda=1} \frac{\partial(\lambda V)}{\partial \lambda} + \left(\frac{\partial U(\lambda S,\lambda V,\lambda n)}{\partial \lambda S}\right)_{\lambda=1} \frac{\partial(\lambda n)}{\partial \lambda} = U$$

$$\frac{\partial U(S,V,n)}{\partial S} S + \frac{\partial U(S,V,n)}{\partial V} V + \frac{\partial U(S,V,n)}{\partial n} n = U,$$

$$TS - PV + \mu n = U,$$

de donde

$$G = \mu n. \tag{A29}$$

Por último, si en la ecuación (A8) de la energía U se reemplaza tanto la entropía S por la temperatura T como el número de moles *n* por el potencial químico μ bajo la transformada de Legendre, se obtiene el *gran potencial termodinámico*

$$\Omega = F - \mu n, \quad donde \quad \Omega = \Omega(T,V,\mu). \tag{A30}$$

En este caso la diferencial de la ecuación (A30) está dada por

$$d\Omega = \left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} dT + \left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} dV + \left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} d\mu . \tag{A31}$$

Por otra parte, la ecuación (A7) se puede expresar de la siguiente manera:

$$dU - SdT + SdT = TdS - PdV + \mu dn - nd\mu + nd\mu , \tag{A32}$$

acomodando términos adecuadamente, esto se reduce a

$$d\left[U - TS - \mu n\right] = -SdT - PdV - nd\mu . \tag{A33}$$

donde

$$\Omega = U - TS - \mu n , \tag{A34}$$

por lo cual

$$\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} = -S , \tag{A35}$$

$$\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} = -P , \tag{A36}$$

$$\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = -n . \tag{A37}$$

En este caso, nótese que de las ecuaciones (A29) y (A30) se obtiene

$$\Omega = F - \mu n = -PV . \tag{A38}$$

Si tuviéramos el sistema en dos fases, con interfaz de separación, entonces

$$\Omega = -PV + \gamma A , \tag{A39}$$

en donde γ es la tensión interfacial y A es el área de la interfaz. Una demostración de la ecuación (A39) se puede obtener considerando: $U = U\left[S, V, A, n\right]$, $F = F\left(T, V, A, n\right)$, $G = G\left(T, P, \gamma, n\right)$, $H = H\left(S, P, \gamma, n\right)$ y $\Omega = \Omega\left[T, V, A, \mu\right]$ de tal manera que

$$\left(\frac{\partial \Omega}{\partial A}\right)_{T,V,\mu} = \left(\frac{\partial U}{\partial A}\right)_{S,V,n} = \left(\frac{\partial F}{\partial A}\right)_{S,V,n} = \gamma , \quad \text{y} \quad \left(\frac{\partial H}{\partial \gamma}\right)_{S,V,n} = \left(\frac{\partial G}{\partial \gamma}\right)_{T,P,n} = -A . \tag{A40}$$

Además, cuando se efectúa trabajo sobre un sistema en equilibrio:

$$dW = -PdV + \gamma dA , \tag{A41}$$

entonces el cambio en la energía queda descrito como

$$dU = TdS - PdV + \gamma dA + \mu dn , \tag{A42}$$

sustituyendo está en la diferencial de la ecuación (A19) se determina que el cambio de la energía libre de Helmholtz esta dado por

$$dF = -SdT - PdV + \gamma dA + \mu dN, \tag{A43}$$

definiendo ahora a la entalpía como H = $U + PV - \gamma A$ se encuentra que

$$dH = TdS + VdP - Ad\gamma + \mu dn, \tag{A44}$$

y definiendo a la energía libre de Gibbs como $G = U + PV - \gamma A - TS$ *se* encuentra que

$$dG = -SdT + VdP - Ad\gamma + \mu dN, \tag{A45}$$

de la definición de la propiedad homogénea de primer orden se tiene que para cualquier λ

$$G(T, P, \gamma, \lambda n) = \lambda G(T, P, \gamma, n), \tag{A46}$$

y derivando con respecto a λ

$$G = \left(\frac{\partial G}{\partial \lambda n}\right)_{\lambda=1,T,P,\gamma} \left(\frac{\partial \lambda n}{\partial \lambda}\right) = n\left(\frac{\partial G}{\partial n}\right)_{T,P,\gamma} = \mu n, \tag{A47}$$

por lo cual

$$U + PV - TS - \gamma A = \mu n, \tag{A48}$$

de la cual por la ecuación (A34) es evidente que se encuentra la ecuación (A39).

Estos son algunos de los conocidos como *potenciales termodinámicos* que son funciones transformadas de Legendre de la energía, y que también constituyen cada uno de ellos una ***R F T***; es decir, de cualquiera de ellos se puede obtener toda la información termodinámica acerca del sistema.

comunicacion-cientifica.com

Anexo B. Energía de interacción de sistemas dipolares

Se sabe que la energía de interacción para sistemas dipolares es

$$W = -\mathbf{p}\cdot\mathbf{E} \tag{B1}$$

y como $\mathbf{F} = -\nabla W$ (B2)

entonces

$$\begin{aligned}\mathbf{F} &= -\nabla\left(-\mathbf{p}\cdot\mathbf{E}\right) = \nabla\left(\mathbf{p}\cdot\mathbf{E}\right)\\ &= \mathbf{E}\times\left(\nabla\times\mathbf{p}\right)+\mathbf{p}\times\left(\nabla\times\mathbf{E}\right)+\left(\mathbf{E}\cdot\nabla\right)\mathbf{p}+\left(\mathbf{p}\cdot\nabla\right)\mathbf{E}\end{aligned} \tag{B3}$$

Como $\mathbf{p}$ es constante, el primer y tercer términos desaparecen, y como $\mathbf{E}$ es conservativo, entonces $\nabla\times\mathbf{E} = 0,$ por lo tanto

$$\mathbf{F} = \left(\mathbf{p}\cdot\nabla\right)\mathbf{E} \tag{B4}$$

Como el trabajo para llevar una partícula desde un punto hasta otro en un campo conservativo está dado por

$$W = -\int \mathbf{F}\cdot d\mathbf{r} \tag{B5}$$

dicho trabajo es la misma energía de interacción. Entonces

$$W = -\int\left(\mathbf{p}\cdot\nabla\right)\mathbf{E}\cdot d\mathbf{r} = -\int\left(\mathbf{p}\cdot\nabla\right)\mathbf{E}\cdot\left(dx\hat{i}+dy\hat{j}+dz\hat{k}\right) \tag{B6}$$

En coordenadas cartesianas

$$\left(\mathbf{p}\cdot\nabla\right)\mathbf{E} = \left(\mathbf{p}\cdot\operatorname{gra} d\right)\mathbf{E} = \left(\mathbf{p}\cdot\operatorname{gra} dE_x\right)\hat{i}+\left(\mathbf{p}\cdot\operatorname{gra} dE_y\right)\hat{j}+\left(\mathbf{p}\cdot\operatorname{gra} dE_z\right)\hat{k} \tag{B7}$$

por lo cual

$$\begin{aligned}W &= -\int\left(\mathbf{p}\cdot\nabla E_x dx+\mathbf{p}\cdot\nabla E_y dy+\mathbf{p}\cdot\nabla E_z dz\right)\\ &= -\int\left\{\mathbf{p}\cdot\left(\frac{\partial E_x}{\partial x}\hat{i}\right)dx+\mathbf{p}\cdot\left(\frac{\partial E_y}{\partial y}\hat{j}\right)dy+\mathbf{p}\cdot\left(\frac{\partial E_z}{\partial z}\hat{k}\right)dz\right\}\\ &= -\int\left(p_x dE_x+p_y dE_y+p_z dE_z\right)\\ &= -\int\mathbf{p}\cdot d\mathbf{E}\end{aligned} \tag{B8}$$

Dado el momento dipolar inducido (ecuación 39), se encuentra que

$$W = -\frac{1}{2}\alpha E^2 \tag{B9}$$

comunicacion-cientifica.com

Anexo C. Teoría de perturbaciones estacionarias

Con esta teoría se trata de encontrar los cambios en los niveles de energía discretos y eigenfunciones de un sistema cuando se aplica una perturbación pequeña. Se supone que el hamiltoniano H en la ecuación de onda de Schrödinger se puede escribir como la suma de dos partes. Una de estas partes H_o es de estructura tan simple que su ecuación de Schrödinger puede ser resuelta, y la otra parte H' es tan pequeña que puede ser tratada como una perturbación de H_o.

$$H = H_0 + H'$$
$$H\psi = W\psi \qquad \text{(C1)}$$
$$H_0 u_n = E_n u_n$$

Caso no-degenerado

La suposición de que H' es pequeño sugiere que desarrollemos las eigenfunciones y eigenvalores perturbados como series de potencias en H'. Esto se hace en términos de un parámetro λ tal que las potencias cero, primera, etc., de λ corresponden a los órdenes cero, primero, etc., del cálculo de perturbación. Reemplazando H' por $\lambda H'$ y expresando ψ y W como series de potencias en λ. Suponiendo que estas dos series son analíticas para λ entre cero y uno, se tiene que los diferentes ordenes de la aproximación de perturbación son dados por los coeficientes de las diferentes potencias de λ.

La función de onda y el nivel de energía perturbados son escritos

$$\psi = \psi_0 + \lambda\psi_1 + \lambda^2\psi_2 + \lambda^3\psi_3 + \cdots$$
$$W = W_0 + \lambda W_1 + \lambda^2 W_2 + \lambda^3 W_3 + \cdots \qquad \text{(C2)}$$

y son sustituidos en la ecuación de onda para dar

$$\left(H_0 + \lambda H'\right)\left(\psi_0 + \lambda\psi_1 + \cdots\right) = \left(W_0 + \lambda W_1 + \cdots\right)\left(\psi_0 + \lambda\psi_1 + \cdots\right) \qquad \text{(C3)}$$

bajo la suposición de que la ecuación (25.3) es válida para un rango continuo de λ, se pueden igualar los coeficientes de potencias iguales de λ en ambos lados, para obtener una serie de ecuaciones que representan sucesivamente órdenes superiores de la perturbación.

$$H_0\psi_0 = W_0\psi_0$$

$$H_0\psi_1 + H'\psi_0 = W_0\psi_1 + W_1\psi_0 \qquad \text{(C4)}$$

$$H_0\psi_2 + H'\psi_1 = W_0\psi_2 + W_1\psi_1 + W_2\psi_0, \text{ etc.}$$

La primera de las ecuaciones (C4) significa que ψ_o es cualquiera de las eigenfunciones no-perturbadas, como se esperaba. Por lo tanto, se pone

$$\psi_0 = u_m, \qquad W_0 = E_m \qquad \text{(C5)}$$

este estado u_m es supuesto ser no-degenerado, aunque otras eigenfunciones no-perturbadas pueden ser degeneradas.

Perturbación a primer orden

Está implícito en el presente tratamiento que el estado u_m no-perturbado es uno de un conjunto discreto (aun cuando el conjunto completo de $u's$ puede ser parcialmente continuo), puesto que de otro modo no interesaría sujetarla al cálculo de la energía perturbada. Expandiendo ψ_1 en términos de u_n

$$\psi_1 = \sum_n a_n^{(1)} u_n \qquad \text{(C6)}$$

La substitución de la ecuación (C6) en la segunda de las ecuaciones (C4) da

$$\sum_n a_n^1 H_0 u_n + H'u_m = E_m \sum_n a_n^1 u_n + W_1 u_m$$

Reemplazando $H_o u_n$ por $E_n u_n$ en el primer término, multiplicando por $\tilde{u}_k$, e integrando sobre todo el espacio,

$$\int a_n^{(1)} \tilde{u}_k E_n u_n d\tau + \int \tilde{u}_k H' u_m d\tau = \int a_n^{(1)} \tilde{u}_k E_m u_n d\tau + \int \tilde{u}_k W_1 u_m d\tau$$

y haciendo uso de la ortonormalidad de las $u's$ se tiene que

$$a_k^{(1)} E_n + \int \tilde{u}_k H' u_m d\tau = a_k^{(1)} E_m + W_1 \delta_{km}$$

es decir, reacomodando términos

$$a_k^{(1)}(E_m - E_k) + W_1 \delta_{km} = \int \tilde{u}_k H' u_m d\tau \equiv H'_{km} \qquad \text{(C7)}$$

donde la integral de la derecha es el elemento de matriz km de la energía H' en la representación en la que el hamiltoniano H_o es diagonal $\left(H_{nm} = \int \upsilon_n(\boldsymbol{r}) H \upsilon_m(\boldsymbol{r}) d\boldsymbol{r}\right)$.

Si tomamos $k = m$ en la ecuación (C7), vemos que

$$W_1 = H'_{mm} \tag{C8}$$

que es el valor esperado de H' para el estado m. La ecuación (C7) también da

$$a_k^{(1)} = \frac{H'_{km}}{E_m - E_k} \qquad k \neq m \tag{C9}$$

entonces una solución a primer orden en H', excepto que $a_m^{(1)}$ aún no es determinada; es obtenida debajo de la normalización de ψ.

Perturbación a segundo orden

Las soluciones de segundo orden en H' se encuentran en la tercera de las ecuaciones (C4) en la que se substituye

$$\psi_2 = \sum_n a_n^{(2)} u_n \tag{C10}$$

para dar

$$\sum_n a_n^{(2)} H_0 u_n + H' \sum_n a_n^{(1)} u_n = E_m \sum_n a_n^{(2)} u_n + W_1 \sum_n a_n^{(1)} u_n + W_2 u_m$$

Como antes, se reemplaza Houn por $E_n u_n$ en el primer término, se multiplica por $\tilde{u}_k$ y se integra sobre todo el espacio para obtener

$$a_k^2 \left(E_m - E_k\right) = \sum_n a_n^{(1)} H'_{kn} - W_1 a_k^{(1)} - W_2 \delta_{km} \tag{C11}$$

Si ahora ponemos $k = m$, vemos con la ayuda de (C8) que

$$W_2 = \sum_n{}' a_n^{(1)} H'_{mn} = \sum_n{}' \frac{H'_{mn} H'_{nm}}{E_m - E_n} = \sum_n{}' \frac{\left|H'_{mn}\right|^2}{E_m - E_n} \tag{C12}$$

donde la prima sobre sigma denota la omisión del término $n = m$ de la suma sobre n.

Similarmente, si $k \neq m$ en la ecuación (C11) obtenemos

$$a_k^{(2)} = \sum_n{}' \frac{H'_{kn} H'_{nm}}{\left(E_m - E_k\right)\left(E_m - E_n\right)} - \frac{H'_{km} H'_{mm}}{\left(E_m - E_k\right)^2} + \frac{a_m^{(1)} H'_{km}}{E_m - E_k}, \qquad k \neq m \tag{C13}$$

Entonces tenemos una solución a segundo orden en H′ excepto que tanto $a_m^{(2)}$ como $a_m^{(1)}$ aún no han sido determinadas. Para encontrarlas se debe normalizar.

Normalización de ψ

Puesto que ψ_0 se escoge igual a u_m , ψ ya está normalizada a orden cero. Entonces, si la integral de normalización $\int |\psi|^2 d\tau$ se pone igual a la unidad para todos los órdenes de λ, cuando ψ está dada por (C2), obtenemos

$\int (\psi_0 \overline{\psi}_1 + \overline{\psi}_0 \psi_1) d\tau = 0$ a primer orden

$\int (\psi_0 \overline{\psi}_2 + \overline{\psi}_0 \psi_2 + |\psi_1|^2) d\tau = 0$ a segundo orden

Esto a la vez da

$$a_m^{(1)} + \overline{a}_m^{(1)} = 0, \qquad a_m^{(2)} + \overline{a}_m^{(2)} + \sum_n \left|a_n^{(1)}\right|^2 \equiv 0$$

Las partes reales de $a_m^{(1)}$ y $a_m^{(2)}$ están fijas por estas relaciones, pero sus partes imaginarias no. La elección de las partes imaginarias de estos coeficientes es equivalente a la elección de una nueva fase para ψ en cada orden del cálculo; esto otra vez afecta la fase del siguiente término de orden superior en ψ. No hay pérdida de generalidad involucrada en hacer la simple elección de cero para estas partes imaginarias, caso en el cual

$$a_m^{(1)} = 0, \qquad a_m^{(2)} = -\frac{1}{2}\sum_n \left|a_n^{(1)}\right|^2$$

Nótese que las energías perturbadas son independientes de esta elección.

La energía y la función de onda a segundo orden en H′ son entonces (poniendo λ=1)

$$W = E_m + H'_{mm} + \sum_n{}' \frac{|H'_{mn}|^2}{E_m - E_n} \tag{C15}$$

$$\psi = u_m + \sum_k{}' \frac{H'_{km} u_k}{E_m - E_k} + \sum_k \left\{ \left[\sum_n{}' \frac{H'_{kn} H'_{nm}}{(E_m - E_k)(E_m - E_n)} - \frac{H'_{km} H'_{mm}}{(E_m - E_k)^2} \right] u_k - \frac{1}{2} \frac{|H'_{km}|^2}{(E_m - E_k)^2} u_m \right\}$$

Se sigue de las ecuaciones (C8) y C(12) que el cálculo de W a un orden dado en H' requiere el conocimiento de ψ solamente al siguiente orden más bajo.

comunicacion-cientifica.com

Anexo D. Átomo de hidrógeno

Para encontrar una expresión de la función de onda que describa el movimiento del núcleo y del electrón de un átomo de hidrógeno, se necesita resolver la ecuación de onda de Schrödinger de la mecánica cuántica. Es importante comentar que la mecánica clásica no puede describir el movimiento de los átomos; sin embargo, si el sistema es suficientemente grande y si nuestra demanda de la medición de exactitud no es demasiado rígida, la mecánica clásica debería proporcionar una buena aproximación al movimiento del sistema. Esto es bien conocido como el límite clásico de la mecánica cuántica.

La ecuación de onda de Schrödinger unidimensional se puede expresar como

$$H\psi(x,t) = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t} \tag{D1}$$

donde ψ es la función de onda que describe el movimiento del átomo y H es un operador cuántico hermitiano tal que el principio de correspondencia se satisface cuando dicho operador H —como función de las variables cuánticas p (momento lineal) y x (posición)— es el mismo que el hamiltoniano clásico como función de las variables clásicas correspondientes. De esta manera, en realidad, H es el operador hamiltoniano:

$$H = \frac{p^2}{2m} + V(x) \tag{D2}$$

es decir, H es la suma de las energías cinética $\left(T = \frac{p^2}{2m}\right)$ y potencial $(V(x))$.

Por lo tanto, el hamiltoniano del átomo de hidrógeno en tres dimensiones se puede expresar como aquel de dos partículas de masas m_1 y m_2:

$$H = -\frac{\hbar^2}{2m_1}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}\right) - \frac{\hbar^2}{2m_2}\left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2}\right) + V(x_1, y_1, z_1, x_2, y_2, z_2, t) \tag{D3}$$

Si ahora la energía potencial depende solo de las coordenadas relativas, tal que $V = V(x_1 - x_2, y_1 - y_2, z_1 - z_2)$. Además, si se definen las coordenadas relativas *x, y, z,* y las coordenadas del centro de masa *X, Y, Z* por

$$\begin{aligned} x &= x_1 - x_2 & y &= y_1 - y_2 & z &= z_1 - z_2 \\ MX &= m_1x_1 + m_2x_2 & MY &= m_1y_1 + m_2y_2 & MZ &= m_1z_1 + m_2z_2 \end{aligned} \tag{D4}$$

Aquí $M = m_1 + m_2$ es la masa total del sistema. En términos de las nuevas coordenadas el hamiltoniano está dado por

$$H = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}+\frac{\partial^2}{\partial Z^2}\right)-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V(x,y,z) \tag{D5}$$

donde $\mu = \dfrac{m_1 m_2}{m_1 + m_2}$ es la masa reducida.

La ecuación de onda de Schrödinger tridimensional se puede resolver por separación de variables:

$$\psi(x,y,z,X,Y,Z,t) = u(x,y,z)U(X,Y,Z)f(t) \tag{D6}$$

con lo cual la ecuación de Schrödinger se reduce a las tres ecuaciones diferenciales

$$-\frac{\hbar^2}{2\mu}\nabla^2 u + Vu = Eu \tag{D7a}$$

$$-\frac{\hbar^2}{2M}\nabla^2 U = E'U \tag{D7b}$$

$$E' + E = -\frac{\hbar}{i}\frac{1}{f}\frac{\partial f}{\partial t} \tag{D7c}$$

donde los operadores laplacianos ∇^2 en la primera y segunda ecuaciones implican diferenciación con respecto a las coordenadas relativas y del centro de masa, respectivamente.

La solución de la ecuación diferencial (D7c) que depende del tiempo es

$$f(t) = Ce^{-\frac{i(E+E')t}{\hbar}} \tag{D8}$$

La ecuación (D7a) describe el movimiento de las dos partículas, y es la misma que la ecuación para el movimiento de una partícula que tiene la masa reducida μ en una energía potencial externa V. La ecuación (D7b) dice que el centro de masa del sistema de dos partículas se mueve como una partícula libre de masa M. Expresando la ecuación (D7b) de la forma

$$\nabla^2 U + \frac{2ME'}{\hbar^2} U = 0 \tag{D9}$$

Notando que para una partícula libre $E' = \frac{p^2}{2m}$ y sabiendo que $\mathbf{p} = \hbar\mathbf{k}$ se tiene que

$$\nabla^2 U + k^2 U = 0 \tag{D10}$$

separando variables se encuentran tres ecuaciones diferenciales ordinarias, una para cada eje de coordenadas cuya solución es (aparte de una constante multiplicativa)

$$U = e^{i\mathbf{k}\cdot\mathbf{r}} \tag{D11}$$

para $\mathbf{k}\cdot\mathbf{r} = k_x x + k_y y + k_z z$ y $k_x^2 + k_y^2 + k_z^2 = k^2$.

Para resolver la ecuación (D7a) es conveniente hacer la transformación de coordenadas rectangulares a coordenadas esféricas, esto se debe a la simetría esférica de la energía potencial en el átomo de hidrógeno. El operador laplaciano en coordenadas esféricas está dado por

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \operatorname{sen}\theta}\frac{\partial}{\partial \theta}\left(\operatorname{sen}\theta \frac{\partial}{\partial \theta}\right) + \frac{1}{r^2 \operatorname{sen}^2\theta}\frac{\partial^2}{\partial \phi^2} \tag{D12}$$

con lo cual la separación de variables debe hacerse para las partes radial y angular de la ecuación (D7a), esto es

$$u(r,\theta,\phi) = R(r)Y(\theta,\phi) \tag{D13}$$

Al dividir entre *u* la ecuación (D7a) se obtiene

$$\frac{1}{R}\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right) + \frac{2\mu r^2}{\hbar^2}\left[E - V(r)\right] = -\frac{1}{Y}\left[\frac{1}{\operatorname{sen}\theta}\frac{\partial}{\partial \theta}\left(\operatorname{sen}\theta \frac{\partial Y}{\partial \theta}\right) + \frac{1}{\operatorname{sen}^2\theta}\frac{\partial^2 Y}{\partial \phi^2}\right] \tag{D14}$$

e igualando a una constante de separación λ se obtienen las ecuaciones diferenciales de las partes radial y angular,

$$\frac{1}{R}\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right) + \frac{2\mu r^2}{\hbar^2}\left[E - V(r)\right] = \lambda \tag{D15}$$

$$-\frac{1}{Y}\left[\frac{1}{\operatorname{sen}\theta}\frac{\partial}{\partial \theta}\left(\operatorname{sen}\theta \frac{\partial Y}{\partial \theta}\right) + \frac{1}{\operatorname{sen}^2\theta}\frac{\partial^2 Y}{\partial \phi^2}\right] = \lambda \tag{D16}$$

La ecuación (D16) se puede resolver por separación de variables

$$Y(\theta,\phi)=\Theta(\theta)\Phi(\phi) \tag{D17}$$

al separar variables se obtienen las ecuaciones diferenciales

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\Phi^2}=-m^2 \tag{D18}$$

y $$\frac{1}{\Theta}\left[\operatorname{sen}\theta\frac{d}{d\theta}\left(\operatorname{sen}\theta\frac{d\Theta}{d\theta}\right)+\beta\operatorname{sen}^2\theta\Theta\right]=m^2 \tag{D19}$$

donde *m* es otra constante usada para separar variables.

De la teoría de momento angular clásica

$$\mathbf{L}=\mathbf{r}\times\mathbf{p}\,. \tag{D20}$$

Cuánticamente, **L** se considera como la misma función de las variables dinámicas cuánticas y, por lo tanto, es un operador vectorial. En el espacio de configuración sus componentes rectangulares son

$$L_x=\left(yp_z-zp_y\right)=\frac{\hbar}{i}\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right) \tag{D21}$$

$$L_y=\left(zp_x-xp_z\right)=\frac{\hbar}{i}\left(z\frac{\partial}{\partial x}-x\frac{\partial}{\partial z}\right) \tag{D22}$$

$$L_z=\left(xp_y-yp_x\right)=\frac{\hbar}{i}\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right) \tag{D23}$$

En coordenadas esféricas estas componentes resultan ser

$$L_x=-\frac{\hbar}{i}\left(\operatorname{sen}\phi\frac{\partial}{\partial\phi}+ctn\theta\cos\phi\frac{\partial}{\partial\phi}\right) \tag{D24}$$

$$L_y=\frac{\hbar}{i}\left(\cos\phi\frac{\partial}{\partial\theta}-ctn\theta\cos\phi\frac{\partial}{\partial\phi}\right) \tag{D25}$$

$$L_z=\frac{\hbar}{i}\frac{\partial}{\partial\phi} \tag{D26}$$

y, por lo tanto

$$L^2=L_x^2+L_y^2+L_z^2=-\hbar^2\left[\frac{1}{\operatorname{sen}\theta}\frac{\partial}{\partial\theta}\left(\operatorname{sen}\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\operatorname{sen}^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \tag{D27}$$

Entonces, la ecuación (D16) es equivalente a

$$L^2Y=\beta\hbar^2Y \tag{D28}$$

para la cual Y define un estado cuya magnitud del momento angular tiene un valor definido igual a $\sqrt{\beta}\hbar$. La ecuación (D18) es equivalente

$$L_z^2\Phi = m^2\hbar^2\Phi \tag{D29}$$

de la cual $$L_z\Phi = m\hbar\Phi \tag{D30}$$

lo cual significa que $Y = \Theta\Phi$ también es una eigenfunción de L_z con eigenvalor $m\hbar$, puesto que las Θ's son funciones de θ solamente.

De manera que la parte angular se ha reducido a un problema de eigenfunciones y eigenvalores del momento angular.

La solución de la ecuación (D18) es, aparte de una constante multiplicativa, igual a

$$\Phi = e^{im\phi} \tag{D31}$$

de manera que la solución general tiene la forma

$$Y(\theta,\phi) = \Theta(\theta)e^{im\phi} \tag{D32}$$

Para que esta función sea univaluada, se requiere que su valor sea el mismo para un punto en la dirección ϕ o en la dirección ϕ+2π; esto significa que debe cumplirse la condición

$$e^{im\phi} = e^{im(\phi+2\pi)} \tag{D33}$$

es decir $$e^{2\pi im} = 1 \tag{D34}$$

lo que demanda que $$|m| = 0,1,2,... \tag{D35}$$

por tanto, la condición de que la función de onda sea univaluada respecto al ángulo azimutal, implica que los eigenvalores de L_z, $\hbar m$, son múltiplos enteros positivos o negativos de $\hbar$: $0, \pm\hbar, \pm 2\hbar, ...$ Puesto que la dirección de z es arbitraria, este resultado debe entenderse en el sentido de que la proyección del momento orbital en *cualquier* dirección está cuantizada; y a *m* se le conoce como el número cuántico *m*.

Entonces, las funciones *Y* son eigenestados simultáneos de L^2 y L_z. Cada uno tendrá los eigenvalores $\beta\hbar^2$ y $m\hbar$ asociados con L^2 y L_z respectivamente. Se puede enfatizar esto poniendo subíndices a las *Y*'s y escribiendo las ecuaciones de eigenvalores L^2 y L_z como

$$L^2 Y_{\beta,m} = \beta\hbar^2 Y_{\beta,m} \tag{D36a}$$

$$L_z Y_{\beta,m} = m\hbar Y_{\beta,m} \tag{D36b}$$

Las ecuaciones (D36) son equivalentes a la ecuación (D16) que, a su vez, es equivalente a las ecuaciones (D18) y (D19). El problema en esta ecuación es encontrar la eigenfunción Θ y los eigenvalores β. La ecuación (D19) es bien conocida en la literatura matemática como la ecuación para las *funciones asociadas de Legendre*. Un método de solución es substituir *x* por $\cos\theta$ en la ecuación (D19) y buscar una solución en series de potencias de *x*. La convergencia de las series en $x = \pm 1$ $(\theta = \pm\pi)$ requiere la terminación de la serie y da condiciones acerca de los eigenvalores de β. Otro método de resolver este problema es aplicando la teoría de momentos angulares con operadores de ascenso y descenso. Por cualquiera de estos métodos se encuentra que

$$\beta = \ell(\ell+1) \tag{D37}$$

donde ℓ puede ser el máximo valor de *m* o también el mínimo valor de *m* para una β dada. Esto es, aplicando L_z a la ecuación (D36b) y restándola de la ecuación (D36a) se tiene

$$\left(L^2 - L_z^2\right)Y_{\beta,m} = \left(\beta - m^2\right)\hbar^2 Y_{\beta,m} \tag{D38}$$

pero $L^2 - L_z^2 = L_x^2 + L_y^2$. Puesto que los eigenvalores de $L_x^2 + L_y^2$ deben ser no-negativos, se puede escribir

$$\left(\beta - m^2\right) \geq 0 \quad \text{así} \quad -\sqrt{\beta} \leq m \leq \sqrt{\beta} \tag{D39}$$

entonces, la serie completa de valores de *m* que corresponden al valor de β dado por la ecuación (D39) son $\ell, \ell-1, \ldots, -(\ell-1), -\ell$.

Usando los polinomios de Legendre se encuentra que la solución de la parte angular son los armónicos esféricos

$$Y_{\ell m}(\theta,\phi) = \left[\frac{2\ell+1}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}\right]^{\frac{1}{2}} P_\ell^m(\cos\theta)e^{im\phi} \tag{D40}$$

La ecuación radial que corresponde al número cuántico ℓ del momento angular es entonces consistente con la ecuación (D15)

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\frac{2\mu}{\hbar^2}\frac{Ze^2}{r}R-\frac{\ell(\ell+1)}{r^2}R=-\frac{2\mu}{\hbar^2}ER \tag{D41}$$

para el potencial $V=-\frac{Ze^2}{r}$, donde Z es el número atómico. Definiendo $\rho=\frac{r}{a}$ se tiene que

$$\frac{1}{a^2}\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{dR}{d\rho}\right)-\frac{\ell(\ell+1)}{a^2\rho^2}R+\frac{2\mu}{\hbar^2}\frac{Ze^2}{a\rho}R+\frac{2\mu E}{\hbar^2}R=0, \tag{D42}$$

y tomando $a^2=-\frac{\hbar^2}{8\mu E}$ y $\lambda=\frac{2\mu Ze^2a}{\hbar^2}$ (D43)

se obtiene

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{dR}{d\rho}\right)+\left(\frac{\lambda}{\rho}-\frac{1}{4}-\frac{\ell(\ell+1)}{\rho^2}\right)R=0. \tag{D44}$$

Eligiendo $\rho R=P(\rho)$ se obtiene

$$\frac{d^2P(\rho)}{d\rho^2}+\left\{\frac{\lambda}{\rho}-\frac{1}{4}-\frac{\ell(\ell+1)}{\rho^2}\right\}P(\rho)=0 \tag{D45}$$

La solución de la ecuación (D44), para la cual *R* es finita conforme $r\to\infty$ y que se comporta apropiadamente en el origen, es decir $rR\to\infty$ conforme $r\to 0$, es

$$R(\rho)=e^{-\frac{\rho}{2}}\rho^{\ell}F(\rho) \tag{D46}$$

donde *F(ρ)* es un polinomio que está relacionado a los *polinomios asociados de Laguerre*. Se escribe así la ecuación (D46) para hacer explícitas las formas asintóticas $e^{-\frac{\rho}{2}}$ conforme $\rho\to\infty$, y ρ^{ℓ} conforme $\rho\to 0$. La condición de que esta serie de potencias en ρ debería terminar es que λ es un entero

$$\lambda=n \qquad n\geq\ell+1 \tag{D47}$$

o $n=\ell+1+n'$ con $n'=0,1,2,...$

Otra vez se ha introducido un entero matemáticamente.

La solución de la función de onda radial a través de los *polinomios asociados de Laguerre* está dada por

$$R_{n\ell}(r) = -\left\{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-\ell-1)!}{2n[(n+\ell)!]^3}\right\}^{\frac{1}{2}} e^{-\frac{1}{2}\rho}\rho^{\ell} L_{n+\ell}^{2\ell+1}(\rho) \tag{D48}$$

donde $a_0 = \dfrac{\hbar^2}{\mu e^2}$ y $\rho = \dfrac{2Z}{na_0} r$. (D49)

Tanto la función radial *R* como la función angular *Y* deben ser normalizadas a la unidad. Por ejemplo, *R* es ahora etiquetada por dos índices *n* y ℓ, de manera que $R_{n\ell}(r)$ es normalizada de acuerdo con

$$\int_0^{\infty} R_{n\ell}^{*}(r) R_{n\ell}(r) r^2 dr = 1. \tag{D50}$$

así que la función de onda total

$$u_{n\ell m} = R_{n\ell}(r) Y_{\ell m}(\theta, \phi) \tag{D51}$$

también es normalizada. Por ejemplo

$$u_{100}(\mathbf{r}) = \left(\frac{Z}{a_0}\right)^{\frac{3}{2}} 2e^{-\frac{Zr}{a_0}}. \tag{D52}$$

Además de la función de onda del átomo de hidrógeno, de las ecuaciones (D43), (D47) y (D49) se obtiene la cuantización de la energía:

$$E_n = -\frac{Z^2 e^2}{2a_0 n^2}; \tag{D53}$$

es decir, con esta ecuación se obtienen los niveles de energía.

comunicacion-cientifica.com

Anexo E. Expansión multipolar

Sea una molécula que consiste en cargas puntuales q_i situadas momentáneamente en los puntos r_i con respecto a un origen dentro de ella. En un punto exterior, a una distancia R del origen, esta distribución producirá un potencial

$$\varphi = \sum \frac{q_i}{|\mathbf{R} - \mathbf{r}_i|} \tag{E1}$$

la cual se puede desarrollar en serie de Taylor vectorial (o tridimensional).[2] En forma vectorial Lloyd (1988) dice que la expansión múltiple en tensores cartesianos es derivada de la expansión de Taylor de $|\mathbf{r} - \mathbf{r}_i|^{-1}$ en términos de $\mathbf{r}_i$, es decir

$$\varphi(\mathbf{r}) = \sum_i q_i \left\{ \frac{1}{r} + (-\mathbf{r}_i)\cdot\nabla\left(\frac{1}{r}\right) + \frac{1}{2!}(-\mathbf{r}_i)(-\mathbf{r}_i):\nabla\nabla\left(\frac{1}{r}\right) + \frac{1}{3!}(-\mathbf{r}_i)(-\mathbf{r}_i)(-\mathbf{r}_i)\cdot\nabla\nabla\nabla\left(\frac{1}{r}\right) + \cdots \right\}.$$

Notando que en tres dimensiones

$$\nabla\left(\frac{1}{r}\right) = \left(\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}\right)\left(\frac{1}{\sqrt{x^2+y^2+z^2}}\right) = \frac{-\frac{1}{2}2x}{\left(x^2+y^2+z^2\right)^{3/2}}\hat{\imath} - \frac{y}{r^3}\hat{\jmath} - \frac{z}{r^3}\hat{k} = -\frac{\mathbf{r}}{r^3},$$

$$\nabla\nabla\left(\frac{1}{r}\right) = \nabla\left(-\frac{\mathbf{r}}{r^3}\right) = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\left(-\frac{(x,y,z)}{\left(x^2+y^2+z^2\right)^{3/2}}\right)$$

$$= \left(\frac{\partial}{\partial x}\left(\frac{-x}{\left(x^2+y^2+z^2\right)^{3/2}}\right), \frac{\partial}{\partial y}\left(\frac{-y}{\left(x^2+y^2+z^2\right)^{3/2}}\right), \frac{\partial}{\partial z}\left(\frac{-z}{\left(x^2+y^2+z^2\right)^{3/2}}\right)\right)$$

---

[2] Serie de Taylor vectorial:

$$f(\mathbf{r}) = f(\mathbf{r}_0) + \nabla f(\mathbf{r}_0)\cdot(\mathbf{r}-\mathbf{r}_0) + \frac{1}{2!}\nabla_i\nabla_j f(\mathbf{r}_0):(\mathbf{r}-\mathbf{r}_0)_i(\mathbf{r}-\mathbf{r}_0)_j + \cdots$$

la cual en dos dimensiones se expresa como

$$f(x,y) = \sum_{n,m=0}^{\infty} \frac{1}{n!m!}\frac{\partial^{n+m}}{\partial x^n \partial y^m} f(x_0,y_0)(x-x_0)^n(y-y_0)^m$$

$$= f(x_0,y_0) + \frac{\partial}{\partial x}f(x_0,y_0)(x-x_0) + \frac{\partial}{\partial y}f(x_0,y_0)(y-y_0) + \frac{\partial^2}{\partial x\partial y}f(x_0,y_0)(x-x_0)$$

$$+ \frac{1}{2!}\frac{\partial^2}{\partial x^2}f(x_0,y_0)(x-x_0)^2 + \frac{1}{2!}\frac{\partial^2}{\partial y^2}f(x_0,y_0)(y-y_0)^2 + \cdots$$

$$= -\frac{x\,\frac{3}{2}\left(x^2+y^2+z^2\right)^{1/2}2x-\left(x^2+y^2+z^2\right)^{3/2}}{\left(x^2+y^2+z^2\right)^3}-\frac{3y^2r-r^3}{r^6}-\frac{3z^2r-r^3}{r^6}$$

$$= -\frac{1}{r^3}\left(\frac{3x^2}{r^2}-1\right)-\frac{1}{r^3}\left(\frac{3y^2}{r^2}-1\right)-\frac{1}{r^3}\left(\frac{3z^2}{r^2}-1\right)$$

y también notando que para $f(x,y,z)=\dfrac{1}{\sqrt{x^2+y^2+z^2}}$ implica que

$$\frac{\partial f}{\partial y}=\frac{-\frac{1}{2}2y}{\left(x^2+y^2+z^2\right)^{3/2}} \quad \text{de donde} \quad \frac{\partial}{\partial x}\frac{\partial f}{\partial y}=\frac{y\,\frac{3}{2}\left(x^2+y^2+z^2\right)^{1/2}2x}{\left(x^2+y^2+z^2\right)^3}=\frac{3xy}{r^5}$$

entonces, el desarrollo en serie de Taylor del potencial (E1) se puede expresar de la siguiente manera

$$\varphi=\frac{1}{R}\sum_i q_i+\frac{1}{R^2}\left(\frac{X}{R}\sum_i q_i x_i+\frac{Y}{R}\sum_i q_i y_i+\frac{Z}{R}\sum_i q_i z_i\right)+\frac{1}{R^3}\left[\frac{1}{2}\left(\frac{3X^2}{R^2}-1\right)\sum_i q_i x_i^2+\frac{1}{2}\left(\frac{3y^2}{R^2}-1\right)\right.$$

$$\left.\times\sum q_i y_i^2+\frac{1}{2}\left(\frac{3Z^2}{R^2}-1\right)\sum_i q_i z_i^2+\frac{3XY}{R^2}\sum_i q_i x_i y_i+\frac{3XZ}{R^2}\sum_i q_i x_i z_i+\frac{3YZ}{R^2}\sum_i q_i y_i z_i\right]+\frac{1}{R^4}[\cdots]$$

$+\cdots$ (E2)

donde X, Y, Z y x_i, y_i, z_i son las componentes de R en r_i, respectivamente. En notación más compacta $\left(X_1=X,\ X_2=Y,\ X_3=Z\right)$ que emplea el *monopolo*

$$q=\sum_i q_i \tag{E3}$$

el vector *dipolar*

$$\mathbf{p}=\sum_i q\mathbf{r}_i \tag{E4}$$

y el tensor *cuadrupolar*

$$\ddot{\mathbf{Q}}=\sum_i q_i\mathbf{r}_i\mathbf{r}_i, \qquad Q_{12}=\sum q_i x_i y_i \qquad \text{etc.} \tag{E5}$$

se puede escribir la expansión multipolar del potencial como

$$\varphi=q\frac{1}{R}+\sum_\rho p_\rho\frac{\partial}{\partial X_\rho}\left(\frac{1}{R}\right)+\frac{1}{2!}\sum_{\rho\sigma}Q_{\rho\sigma}\frac{\partial}{\partial X_\rho}\frac{\partial}{\partial X_\sigma}\left(\frac{1}{R}\right)+\cdots \tag{E6}$$

Términos superiores en esta serie incluyen multipolos superiores, que son tensores de rango superior construible por una extensión de (E5). Nótese que para una distribución de carga

esférica simétrica, el término proporcional a R^{-3} en φ (el así definido término cuadrupolar) desaparece, aunque por definición (E5) el tensor cuadrupolar sigue finito (Margenau y Kestner, 1971).

Una expansión equivalente bien conocida del potencial φ que incluye los polinomios de Legendre P_n está dada por (Margenau y Murphy, 1956).

$$\varphi(R) = \sum_{n=0}^{\infty} \frac{1}{R^{n+1}} \sum_i q_i r_i^n P_n(\cos\theta_i) \tag{E7}$$

siendo θ_i en ángulo entre $\mathbf{r}_i$ y $\mathbf{R}$.

Si consideramos dos moléculas como en el caso del átomo de hidrógeno, la molécula 1 1 es una distribución de cargas como la considerada anteriormente; mientras que la molécula 2 es una distribución de cargas q_j, dispuesta en puntos $\mathbf{r}_j$. Estas coordenadas posteriores están medidas desde un nuevo origen situado en la molécula 2, y $\mathbf{R}$ es la distancia vectorial desde el origen del primer sistema de coordenadas localizado en 1 aquel del segundo localizado en la molécula 2. Si cada $r_j < R$, la energía potencial entre 1 y 2 es

$$W = \sum_j q_j \varphi\left(\left|\mathbf{R} + \mathbf{r}_j\right|\right) \tag{E8}$$

y desarrollando una vez más en serie de Taylor se tiene que

$$W = \sum_j q_j \varphi + \sum_j q_j \left(x_j \frac{\partial \varphi}{\partial X} + y_j \frac{\partial \varphi}{\partial Y} + z_j \frac{\partial \varphi}{\partial Z} \right) + \frac{1}{2!} \sum_j q_j \left(x_j^2 \frac{\partial^2 \varphi}{\partial X^2} + \cdots + 2 y_j z_j \frac{\partial^2 \varphi}{\partial Y \partial Z} \right) + \cdots$$

$$= \left\{ q' + \sum_\rho p'_\rho \frac{\partial}{\partial X_\rho} + \frac{1}{2!} \sum_{\rho\sigma} Q'_{\rho\sigma} \frac{\partial^2}{\partial X_\rho \partial X_\sigma} + \cdots \right\} \varphi(R). \tag{E9}$$

Las cantidades primadas son momentos de carga para la molécula 2. Una expresión compacta de (E9), que ya no se va a expresar aquí, se obtiene con la ayuda de tensores irreducibles. Los primeros términos de (E9), que son de mayor utilidad en la teoría de fuerzas intermoleculares, toman una forma simple si el eje-z se toma a lo largo de $\mathbf{R}$ para todas las coordenadas (así que X=Y=0, Z=R):

$$W = \frac{qq'}{R} + \frac{q'p_z - qp'_z}{R^2} + \frac{1}{R^3}\left(q'w_3 + qw'_3\right) + \frac{1}{R^4}\left(q'w_4 - qw'_4\right) + \frac{1}{R^5}\left(q'w_5 + qw'_5\right)$$
$$+ \frac{1}{R^3}\sum_{ij} q_i q_j \left(x_i x_j + y_i y_j - 2z_i z_j\right) + \frac{3}{2R^4}\sum_{ij} q_i q_j \left[r_i^2 z_j - z_i r_j^2 + \left(2x_i x_j + 2y_i y_j - 3z_i z_j\right)\left(z_i - z_j\right)\right]$$
$$+ \frac{3}{4R^5}\sum_{ij} q_i q_j \left[r_i^2 r_j^2 - 5z_i^2 r_j^2 - 5r_i^2 z_j^2 - 15z_i^2 z_j^2 + 2\left(4z_i z_j - x_i x_j - y_i y_j\right)^2\right] + \frac{1}{2R^5}\sum_{ij} q_i q_j \left[3\left(r_i^2 + r_j^2\right)\right.$$
$$\left.\times\left(4z_i z_j - x_i x_j - y_i y_j\right) + 5\left(z_i^2 + z_j^2\right)\left(3x_i x_j + 3y_i y_j - 4z_i z_j\right)\right] + O\left(R^{-6}\right) \text{ (eje-Z a lo largo de } \mathbf{R}\text{)} \qquad \text{(E10)}$$

en esta expresión se ha usado

$$w_n \equiv \sum_i q_i r_i^n P_n\left(\cos\theta_i\right)$$

$$w'_n \equiv \sum_i q'_i r_i^n P_n\left(\cos\theta_i\right)$$

Los primeros cinco términos de (E10) representan interacciones monopolo-monopolo, monopolo-dipolo, etc. Los cuales desaparecen si ambas moléculas son neutras. Los términos restantes proporcionales a R-3 y R-4 son conocidas como las energías dipolo-dipolo y dipolo-cuadrupolo; hay dos componentes en R-5, la primera representa la interacción de dos cuadrupolos, la segunda aquella de un dipolo formado entre un octupolo con otro octupolo.

comunicacion-cientifica.com

Anexo F. Gradiente del potencial

La ley de Gauss en la forma de divergencia (ecuación 142), no es suficiente para especificar completamente las tres componentes del campo eléctrico. Se sabe que un campo vectorial se puede especificar completamente si su divergencia y su rotacional son dados en todo el espacio. Sin embargo, si se expresa la ley de Gauss en la forma (Jackson, 1962):

$$\mathbf{E}(\mathbf{x}) = \int \rho(\mathbf{x}') \frac{(\mathbf{x}-\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|^3} d^3x' \tag{F1}$$

el factor vectorial en el integrando, visto como una función de **x**, es el gradiente negativo del escalar $1/|\mathbf{x}-\mathbf{x}'|$:

$$\frac{(\mathbf{x}-\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|^3} = -\nabla\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right) \tag{F2}$$

como el gradiente se aplica a **x**, pero no a la variable de integración **x′**, se puede sacar de la integral. Entonces el campo eléctrico es

$$\mathbf{E}(\mathbf{x}) = -\nabla\left(\int \frac{\rho(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} d^3x'\right) \tag{F3}$$

definiendo el argumento del gradiente como una función escalar conocida como el potencial, entonces se obtiene la ecuación

$$E = -\nabla\Phi \tag{F4}$$

y como el rotacional del gradiente de cualquier función escalar desaparece $\nabla \times \nabla\Phi = 0$, por lo tanto

$$\nabla \times \mathbf{E} = 0 \tag{F5}$$

es la ecuación que se necesita para poder especificar completamente las tres componentes del campo eléctrico. Por otra parte, que el rotacional del campo eléctrico es cero [ecuación (16)] también es evidente si se considera el trabajo para mover una carga puntual a lo largo de una trayectoria cerrada, lo cual implica que

$$\oint_L \mathbf{E} \cdot d\ell = 0$$

ya que empleando el teorema de Stokes se obtiene

$$\oint_L \mathbf{E} \cdot d\ell = \int_S (\nabla \times \mathbf{E}) \cdot d\boldsymbol{S}$$

en consecuencia, se obtiene la ecuación (F5).

comunicacion-cientifica.com

Anexo G. Quaterniones

Un cuaternión **Q** es un conjunto de cuatro cantidades escalares

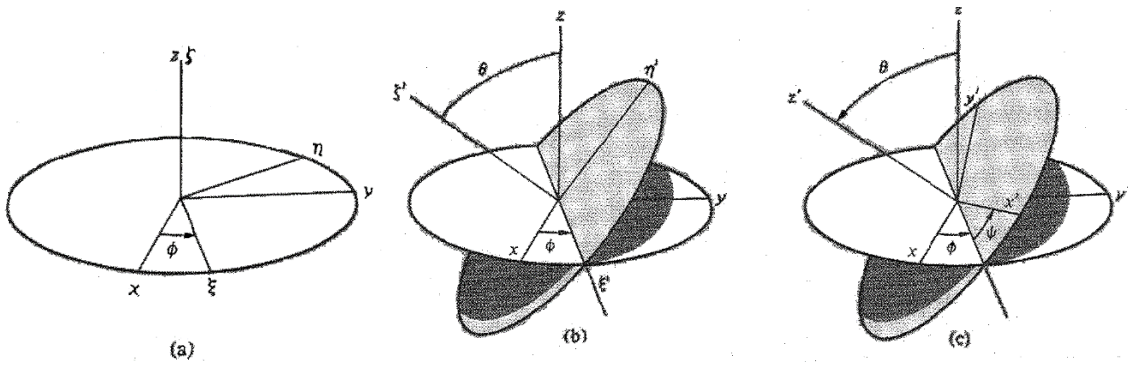
$$\boldsymbol{Q} = (q_0, q_1, q_2, q_3)$$

y frecuentemente se piensa de los últimos tres elementos (q_1, q_2, q_3) como constituyentes de un vector. Aquí los cuaterniones de interés satisfacen la constricción

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$$

la manera en que tal cuaternión puede representar la orientación de un cuerpo rígido es discutida por Goldstein (1980). En la convención de ángulos Euler de la figura 57.

Figura 57. *Ángulos de Euler*



Nota. (a) Sistema inicial xyz rotado un ángulo ϕ alrededor del eje z en contra de las manecillas del reloj, resultando en los ejes ξηζ. (b) Los ejes ξηζ son rotados un ángulo θ alrededor del eje ξ en contra de las manecillas del reloj, resultando en los ejes ξ′η′ζ′. (c) Los ejes ξ′η′ζ′ son rotados un ángulo $\boldsymbol{\psi}$ alrededor del eje ζ′ en contra de las manecillas del reloj, resultando en los ejes deseados x′y′z′.

y la matriz de rotación de los cosenos directores, es más conveniente definir

$$q_0 = cos\frac{1}{2}\theta cos\frac{1}{2}(\theta + \psi)$$

$$q_1 = sen\frac{1}{2}\theta cos\frac{1}{2}(\theta - \psi)$$

$$q_2 = sen\frac{1}{2}\theta sen\frac{1}{2}(\theta - \psi)$$

$$q_3 = cos\frac{1}{2}\theta sen\frac{1}{2}(\theta + \psi)$$

Cuando la matriz de rotación se convierte a

$$A = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}$$

Los cuaterniones de cada molécula satisfacen las ecuaciones de movimiento

$$\begin{pmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix}\begin{pmatrix} 0 \\ \omega_X^b \\ \omega_Y^b \\ \omega_Z^b \end{pmatrix}$$

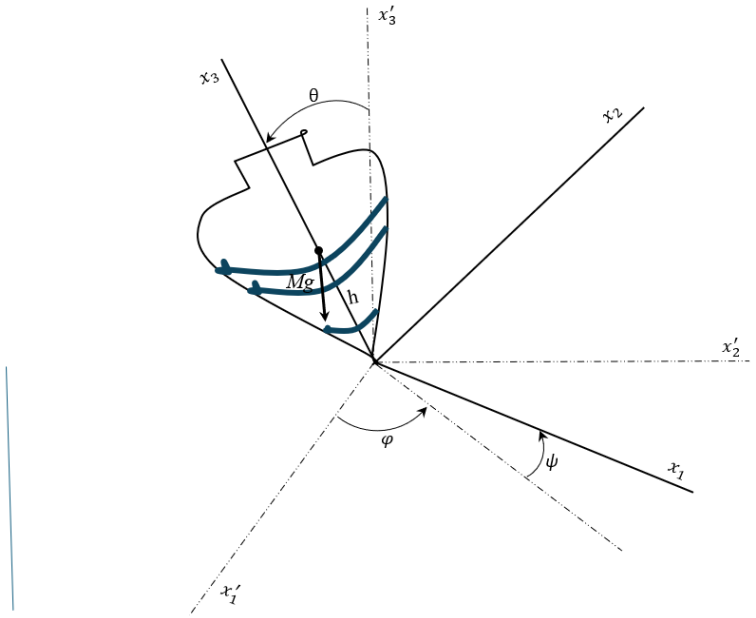
Estas ecuaciones de movimiento con ecuaciones resultantes para las componentes del vector ω (Allen y Tildesley, 1987), y usando la matriz de rotación en términos de los ángulos de Euler para transformar entre coordenadas de espacio-fijo y de cuerpo-fijo no contienen singularidades. Son un sistema de ecuaciones diferenciales que se pueden resolver por el método predictor-corrector.

comunicacion-científica.com

Anexo H. Trompo

Para el movimiento de un trompo consideremos uno simétrico, rotante, con un punto fijo girando en un campo gravitacional como se ve en la figura 58. Su energía cinética traslacional

Figura 58. *Trompo*



Fuente: elaboración propia.

Tomando los cosenos directores de las figuras 41 y 57 en sus matrices B, D, y C ahora como

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta & sen\theta \\ 0 & -sen\theta & cos\theta \end{pmatrix}$$

La matriz A=BCD de los cosenos directores con sus a_{ij} componentes está dada por

$$A = \begin{pmatrix} cos\phi cos\psi - sen\phi cos\theta sen\psi & sen\phi cos\psi + cos\phi cos\theta cos\psi & sen\phi sen\psi \\ -cos\phi sen\psi - sen\phi cos\theta cos\psi & -sen\phi sen\psi + cos\phi cos\theta cos\psi & cos\psi sen\theta \\ sen\theta sen\psi & -sen\theta cos\phi & cos\theta \end{pmatrix}$$

comunicacion-cientifica.com

Referencias

Allen, M. P. y Tildesley, D. J. (1987). *Computer simulation of liquids.* Oxford University Press.

Arfken, G. B. (1985). *Mathematical methods for physicists.* Academic Press.

Atkins, P. W. (2001). *Physical chemistry*. Oxford University Press.

Atkins, P. W. y Friedman, R. S. (1996). *Molecular quantum mechanics*. Oxford University Press.

Atkins, P. W. y Friedman, R. S. (2005). *Molecular quantum mechanics* (4ª ed.). Oxford University Press.

Atkins, P., Jones, L. y Laverman, L. (2008). *Chemical principles he quest for insight*. W. H. Freeman and Company.

Berkowski, K. (2005). *Introduction to Organic Chemistry, Review, Molecular Structure* [curso en línea]. MIT Open Courseware.

Blokhintsev, D. I. (1964). *Quantum mechanics*. D. Reidel Publishing.

Bravo Pérez, G. (1986). *Estructura electrónica y análisis vibracional de los superoxidos* [Tesis de maestría]. UAM.

Callen, H. B. (1996). *Thermodynamics*, J. Wiley.

Castillo, S., Pacheco-Sánchez J. H., Luna-García, H. y Cruz, A. (2005). *Ab initio* study of the interaction of a Pt4 pyramid cluster with the CO molecule. En Dorzatti, P., Groppi, G., Ciambelli, P. y Sannino D (Eds.). *Catalytic Combustion* (vol. 1., pp.118-122). Polipress.

Dassault Systemes Biovia Corp. (2017). *BIOVIA Materials Studio 2017* (versión 17.1.0) [software]. https://biovia-materials-studio-2017.software.informer.com/

Davydov, A. S. (1965). *Quantum mechanics.* Pergamon Press.

Desales-Guzmán, L. A., Pacheco-Sánchez, J.H., Isidro-Ortega, F. J. y García-Rosales, G. (2018). Modelling carbyne C12-ring calcium decorated for hydrogen storage. *Revista Mexicana de Física, 64*(6), 634-641.

Drennan, C. L. y Taylor, E. M.V. (2008). *Principle of Chemical Science* [curso en linea]. MIT Open Courseware. https://ocw.mit.edu/courses/5-111sc-principles-of-chemical-science-fall-2014/

Durand, P. H. y Barthelat, J. C. (1975). A theoretical method to determine atomic pseudopotentials for electronic structure calculations of molecules and solids. *Theoretica Chimica Acta, 38*, 283-302.

Esquivel-Aguilar, Y. C. (2015). *Optimización geométrica por simulación molecular del carbón activado* [Tesis de maestría]. Instituto Tecnológico de Toluca.

Esquivel-Aguilar, Y. C., Pacheco-Sánchez, J. H. y García-Rosales, G. (2019). DFT Study about adsorption of a C14 carbon ring on an organochlorinated DDE compound. *American Journal of Biomedical Science y Research, 2*(3), 98-102. 10.34297/AJBSR.2019.02.000581

Evans, D. J. (1977). On the representatation of orientation space. *Molecular Physics, 34*(2), 317-325.

Flores-Ramírez, L. y Pacheco-Sánchez, J. H. (2019). Antibiotic-receptor molecular interactions. *American Journal of Biomedical Science y Research, 1*(5), 203-213.

Garstang, R. (1966). *The physical constants relating to atomic spectra. Conference paper*. Academic Press.

Geerlings, P., De Proft, F. y Langenaeker, W. (2003). Conceptual density functional theory. *Chemical Reviews*, *103*, 1793-1873.

Giancoli, D. C. (2009). *Physics for scientists & engineers* (vol. 3). Upper Saddle River.

Goldstein, H. (1980). *Classical mechanics*. Addison-Wesley Series in Physics.

Gupta, V. P. (2016). *Principles and applications of quantum chemistry*. Academic Press.

Hernández-Benitez, D. y Pacheco-Sánchez, J. H. (2018). Optimization of chitosan+activated carbon nanocomposite. DFT study. *Archives of Organic and Inorganic Chemical Sciences, 3*(5), 436-448.

Herzberg, G. (1944). *Atomic spectra and atomic structure* (2ª ed.). Dover Publications.

Hirschfelder, J. O., Curtiss, F. Ch. y Bird, R. B. (1954). *Molecular theory of gases and liquids*. John Wiley and Sons.

Hunter, R. J. (1987). *Foundations of colloid science*. Clarendon Press.

Isidro-Ortega, F. J., Pacheco-Sánchez, J. H. y. Desales-Guzmán, L. A. (2017). Hydrogen storage on lithium decorated zeolite templated carbón, DFT study. *International Journal of Hydrogen Energy, 42*(52), 30704-30717. https://doi.org/10.1016/j.ijhydene.2017.10.098.

Isidro-Ortega, F. J., Pacheco-Sánchez, J. H., Alejo, R., Desales-Guzmán, L. A. y Arellano, J. S. (2019) .Theoretical studies in the stability of vacancies in zeolite templated carbon for hydrogen storage. *International Journal of Hydrogen Energy, 44*(13), 6437-6447. https://doi.org/10.1016/j.ijhydene.2019.01.196.

Israelachvili, J. N. (1985). *Intermolecular and surface forces* (3ª ed.). Academic Press.

Israelachvili, J. N. (1992). *Intermolecular and surface forces*. Academic press.

Israelachvili, J. N. y Pashley, R. M. (1982). La interacción hidrofóbica es de largo alcance y decae exponencialmente con la distancia. *Nature*, 300, 341-342.

Jackson, J. D. (1962). *Classical electrodynamics*. John Wiley and Sons.

Landau, L. D. y Lifshitz, E. M. (1958). *Quantum mechanics-non-relativistic theory*. Pergamon.

Landau, L. D. y Lifshitz, E. M. (1963). *Statiscal physics. Second revisen an enlager edition (vol. 5)*. Pergamon Press.

Larrea Munguía, J. M., Pacheco-Sánchez, J. H. y Del Razo López, F. (2017). Adsorption on (Ni-H_2, Pd-H_2, Pt-H_2) metal-hydrogen interaction: DFT approach. *American Journal of Quantum Chemistry and Molecular Spectroscopy, 1*(1), 26-39. 10.11648/j.ajqcms.20160101.14

Leighton, R. B. (1959). *Principles of modern physics*. McGraw-Hill Book Company.

Lloyd L. L. (1988). *Molecular thermodynamics of non-ideal fluids*. Butterworths.

London, F. (1930). Zur theorie und systematik der molekularkrafte. *Zeitschrift für Physik*, *63*, 245-279. http://dx.doi.org/10.1007/BF01421741

London, F. (1937). The general theory of molecular forces. *Transactions od the Faraday Society*, *33*(8), 8b-26.

Löwdin, P. O. (1958). Correlation problem in many-electron quantum mechanics i. review of different approaches and discussion of some current ideas. *Advances in Chemical Physics*, *2*, 207-322.

Mahanty, J. y Ninham, B. W. (1976). *Dispersion forces*. Academic Press.

Margenau, H. y Kestner, N. R. (1971). *Theory of intermolecular forces*. Pergamon.

Margenau, H. y Murphy, G. M.(1956). *Physics and chemistry* (Vol. I). D. van Nostrand Company.

Marion, J. (1965). *Classical electromagnetic radiation* (2ª ed.). Academica Press.

Márquez, M., Carbajal, G. y Pacheco, J. (2016). Estado de transición de la reacción de síntesis de goetita mediante simulación molecular. *Revista de Simulación y Laboratorio, 3*(7), 25-32.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. y Teller, E. (1953). Equation of state claculations by fast computing machines. *The Journal of Chemical Physics*, 21(6).

McMurry, J. (2004) *Organic chemistry*. Thomson Brooks/Cole.

Möller, C. y Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. *Physical Review Journals Archive, 46*(7), 618-622.

Novaro, O., Pacheco-Blas, M. A. y. Pacheco-Sánchez, J. H. (2012). Potential energy surfaces for reactions of x metal atoms (X=Cu, Zn, Cd, Ga, Al, Au, orHg) with YH_4 Molecules (Y = C, Si, or Ge) and transition probabilities at avoided crossings in some cases. *Advances in Physical Chemistry, 2012*(1), 1-17. doi: 10.1155/2012/720197

Oxtoby, D. W., Nachtrieb, N. H. y Freeman, W. A. (1990). *Chemistry*. Saunders GSS.

Pacheco Blas, M. A. (2009). *Probabilidades de transición en la reaciones metal-gas* [Tesis de licenciatura]. Universidad Nacional Autónoma de México.

Pacheco-Sánchez, J. H. (2019a). Some pollutants to defeat to get clean drinking water. *American Journal Biomedical Science & Research, 4*(6), 429-430.

Pacheco-Sánchez J. H. (2019b). Chitosan copolymer with activated carbon as nanocomposite through DFT modeling. *American Journal of Biomedical Science y Research, 4*(6), 403-408. 10.34297/AJBSR.2019.04.000843

Pacheco-Sanchez, J. H. y Mora-Zarco, K. (2017, 10-14 de octubre). Pt(3D,1S)-O2 Interaction by ab initio calculations. *EMN Meeting on Computation and Theory* [Conferencia]. Las Vegas.

Pacheco-Sanchez, J. H. y Mora-Zarco, K. (2019). Preliminary ab initio calculations on Pt(3D,1S)-O2 interaction. *American Journal of Biomedical Science & Research, 6*(6). 10.34297/AJBSR.2019.06.001091

Pacheco-Sánchez, J. H., Castillo, S., Luna-García, H. y Novaro, O. (2007). Transition probabilities on Ga (2P, 2S, and 2P) +CH4 reactions. *Journal of Chemical Physics*, *126*(10).

Pacheco-Sánchez, J. H., Luna-García, H. y Castillo, S. (2004). *Ab initio* study of the reactions of Ga (2P, 2S, and 2P) with methane. *Journal of Chemical Physics, 120*(9), 4240-4246.

Pacheco-Sánchez, J. H., Zaragoza-Rivera, I. P. y Bravo-Ortega, A. (2017). Interaction of small carbon molecules and zinc dichloride: DFT study. *Revista Mexicana de Física, 63*(1), 97-110.

Paglia, S. R. (1971). *Introductory quantum chemistry*. Harper y Row.

Parr, R. G. y Yang, W. (1989). *Density-functional theory of atoms and molecules*. Oxford University Press.

Parr, R. G., Donnelly, R. A., Levy, M. y Palke, W. E. (1978). Electronegativity: the density functional viewpoint. *The Journal of Chemical Physics, 68*, 3801-3807. https://doi.org/10.1063/1.436185

Pauling, L. y Beach, J. (1935). The van der waals interaction of hydrogen atoms. *Physical Review Journals Archive, 47*(686). https://doi.org/10.1103/PhysRev.47.686

Phillips, P. y Davidson, E. R. (1984). The chemical potential for interacting fermions in a harmonic potential. En Dahl, J. P y Avery, J. (Eds.). *Local density approximations in quantum chemistry and solid state physics* (pp. 43-52). Plenum.

Pople, J. A., Binkley, J. S. y Seeger, R. (1976.). Theoretical models incorporating electron correlation. *International Journal of Quantum Chemistry, 10*(10), 1-19.

Prigogine, I. y Defay, R. (1954). *Chemical thermodynamics* (1ª Ed.). Longmans Green and Co. London.

Richards, W. G. y Cooper, D. L. (1983). *Ab initio molecular orbital calculations for chemists*. Oxford Science Publications.

Rivas Castro, S. F. (2016). *Modelación del carbón activado mediante KOH e identificación de la estructura molecular en anillos* [Tesis de maestría]. Instituto Tecnológico de Toluca.

Ross, R. B., Powers, J. M., Atashroo, T., Ermler, W. C., LaJohn, L. A., Christiansen, P. A. (1990). Ab initio relativistic effective potentials with spin-orbit operators. IV. Cs through Rn. *The Journal of Chemical Physsics, 93*(9), 6654-6670.

Ryckaert, J. P. y Bellemans, A. (1978). Molecular dynamics of liquid alkanes. *Faraday Discussions of the Chemical Society*, *66*, 95-106.

Ryckaert, J. P., Ciccotti, G. y Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics, 23*(3), 327-341.

Saxon, D. S. (1968). *Elementos de mecánica cuántica*. EASO.

Schiff, L. I. (1955). *Quantum mechanics*. McGraw-Hill.

Slater, J. C. y Kirkwood, J. (1931). The Van Der Waals Forces in Gases. *Physical Review Journals Archive*, 37(682). https://doi.org/10.1103/PhysRev.37.682

Sóbol, I. M. (1976). *Método de Montecarlo*. MIR.

Suárez, R. y Pacheco, J. (2017). Estudio DFT: adsorción de cobre doblemente ionizado por el grupo amino del quitosano. *Revista de Simulación y Laboratorio, 4*(13), 46-52.

Suárez-Reyes, R., Zaragoza, P. I. y Pacheco Sánchez, H. J. (2018). BOMD study on the reaction among one amaranth and one chitosan repetitive unit. *Mediterranean Journal of Chemistry, 7*(4), 243-252.

Sun, H. (1998). The COMPASS force field: parameterization and validation for phosphazenes. *Computational and Theoretical Polymer Science, 8*(l), 229-246.

Szabo, A. y Ostlund, N. S. (1996). *Modern quantum chemistry*. Dover Publications.

Valenzuela-Hermocillo, E., Pacheco-Sánchez, J. H. y Rivas-Catro, S. (2019). Modelado molecular de un ánodo de carbón activado como soporte de platino por medio de cálculos DFT. *Revista de Energía Química y Física, 6*(19), 31-38. 10.35429/JCPE.2019.19.6.31.38

Zaragoza, I. P., Pacheco-Sánchez, J. H., Echevarria-Chan, I. y Bravo-Ortega, A. (2014). DFT study of interaction between a hydrogen molecule and AgY-zeolite. *Revista Mexicana de Física, 60*(6), 460-465.

Zaragoza, I. P., Santamaria, R., Bokimi, X. y Suarez, M. V. E. (2018). A DFT study of the interaction of bimetallic Pt-Sn with Ethanol. *Mediterranean Journal of Chemistry, 6*(6), 231-237.

Zumdahl, S. S. y Decoste, D. J. (2014). *Basic chemistry*. Houghton Mifflin Company.

comunicacion-cientifica.com

Sobre el autor

Juan Horacio Pacheco Sánchez

Es Doctor en Ciencias e Ingeniería de Materiales por la Universidad Autónoma Metropolitana (UAM), Maestro en ciencias en Física por la Universidad Autónoma de México (UNAM) y Físico también por la UNAM. Se desempeña como profesor de Física, Matemáticas e Inteligencia Artificial en el TecNM / Instituto Tecnológico de Toluca desde 2006. Tiene nivel 2 en el SNII. Es miembro activo de la Sociedad Mexicana de Física y revisor de revistas en las compañías: ACS, RSC, Elsevier, Wiley y SMF. Forma parte de la Red de Producción y Almacenamiento de Energía (RedPAE), de la Red Mexicana de Fisicoquímica Teórica (RMFQT) y de la Red de Nanociencia. Su principal área de interés es la fisicoquímica del carbono mediante ingeniería molecular, particularmente en energía y medioambiente.

ORCID: https://orcid.org/0000-0002-9525-5473

Scopus: 56029170300

Correo electrónico: jpachecos@toluca.tecnm.mx

Fundamentos de física en ingeniería molecular, de Juan Horacio Pacheco Sánchez, publicado por Ediciones Comunicación Científica, S. A. de C. V., se terminó de imprimir en mayo de 2026, en Litográfica Ingramex, S. A. de C. V., Centeno 162-1, Granjas Esmeralda, 09810, Ciudad de México. El tiraje fue de 50 ejemplares impresos y en versión digital para acceso abierto en los formatos PDF, Epub y HTML. El cuidado de la edición estuvo a cargo de Fidel Carlón Solís

Este libro pone énfasis en la física elemental con la que se construye la ingeniería molecular. En el capítulo uno se aplican tanto la mecánica cuántica como las fuerzas intermoleculares para explicar la construcción de la estructura molecular de sistemas basados en interacciones *núcleo-electrón* y *electrón-electrón* para una energía cinética electrónica.

En el capítulo dos se emplea el cálculo variacional para resolver las ecuaciones de Hartree-Fock a partir tanto del método SCF de Roothaan como de funciones base pertinentes. Esto es lo mínimo necesario para introducir la teoría de funcionales de la densidad (DFT). También se introducen los métodos semiempíricos, con el hamiltoniano ϖ-electrónico, efectivo con aplicación en sistemas hasta diez veces mayores que DFT; así como la mecánica molecular basada en mecánica clásica, con aplicación en sistemas mucho mayores que los métodos semiempíricos y tan grandes como los recursos computacionales lo permitan.

En el capítulo tres se explica cómo la dinámica molecular tiene esencia en las ecuaciones de movimiento de sistemas atómicos, que se resuelven numéricamente utilizando, por ejemplo, el método de Verlet o algún otro análogo.

La parte esencial del capítulo cuatro destaca el método de Montecarlo como un método alternativo a la dinámica molecular. De manera sencilla, se explica cómo es que el método de Montecarlo requiere de bases de datos muy grandes para tener buenas aproximaciones a los resultados.

Finalmente, el quinto y último capítulo es una introducción a la aplicación de *software* de simulación.



Juan Horacio Pacheco Sánchez es doctor en Ciencias e Ingeniería (de Materiales) por la UAM-A, Maestro en Ciencias (Física) y Físico por la FC-UNAM. Se desempeña como profesor-investigador en TecNM/Instituto Tecnológico de Toluca, es nivel II en SNII y forma parte de las redes-investigación: RMFQT, Nanociencias y RedPAE.

RENIECYT
Registro Nacional de Instituciones y Empresas Científicas y Tecnológicas
2000922

DOI.ORG/10.52501/CC.421

ISBN: 978-968-9738-97-8

EDICIONES COMUNICACIÓN CIENTÍFICA PUBLICACIONES ARBITRADAS
comunicacion-cientifica.com