



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

Handwritten notes and stamps

DINAMICA MOLECULAR APLICADA A
FRONTERAS DE GRANO.

T E S I S

QUE PARA OBTENER EL TITULO DE

F I S I C O

P R E S E N T A:

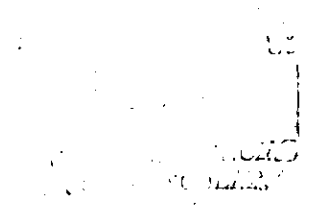
IAN ALEJANDRO SIGAL SEFCHOVICH

DIRECTOR DE TESIS: DR. LUIS DAVID ROMEU CASAJUANA

MEXICO, D. F.

281452

2000





Universidad Nacional
Autónoma de México

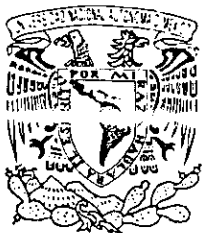


UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

MAT. MARGARITA ELVIRA CHÁVEZ CANO
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis: Dinámica Molecular aplicada a fronteras de grano.

realizado por IAN ALEJANDRO SIGAL SEFCHOVICH

con número de cuenta 9350501-9, pasante de la carrera de FÍSICA

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis

Propietario DR. LUIS DAVID ROMEU CASAJUANA

Propietario DR. IGNACIO LUIS GARZON SOSA

Propietario DRA. MARCELA REGINA BELTRAN SANCHEZ

Suplente DR. ALIPIO GUSTAVO CALLES MARTINEZ

Suplente M. EN C. LUIS MANUEL BELTRAN DEL RIO CABALLERO

Consejo Departamental de FÍSICA
DRA. PATRICIA GOLDSTEIN MEVACHE
 FACULTAD DE CIENCIAS
 DEPARTAMENTO DE FÍSICA



¿Se la dedico a mi mama o a toda mi familia?

Si

Gracias David por confiar en mi y ayudarme a vivir mas contento.

Gracias mami por tanto.

Gracias Carlos, Rodrigo, Ek, Sol, Jimena, Isaac, Shu, Memo, Aida,
hermanastros y hermanastras.

Gracias Leandro, Fidel, Adolfo y Carlitos.

Gracias a Luis Beltran, Ignacio Garzon, Marcela Beltran, Alipio Calles y
Alfredo Gomez.

Gracias a mis maestros y amigos de la Facultad de Ciencias que me
mostraron un mundo infinitamente mas rico del que yo imaginaba.

Índice General

0.1	Prólogo	xii
0.1.1	Propósitos	xii
0.1.2	Una visión panorámica	xii
I	El problema	1
1	Fronteras de grano	3
1.1	Importancia científica y tecnológica	3
1.2	Clasificación	3
1.3	Modelos	4
1.3.1	Fronteras de bajo ángulo:	5
1.3.2	Fronteras de alto ángulo:	6
1.3.3	Modelo de estructura amorfa:	6
1.3.4	Modelo de Islas	6
1.3.5	Modelo de sitios de coincidencia (CSL)	7
1.3.6	Modelo de Casi Coincidencia	11
1.3.7	Modelo de Planos Coincidentes	11
1.3.8	Modelo de burbujas.	12
1.3.9	Modelo de la "O-Lattice"	13
1.3.10	Red DSC	15
1.4	Características	15
1.4.1	Orientación de los granos	16
1.4.2	Fronteras especiales	16
1.4.3	Propiedades	16
1.5	Crítica de los modelos	21
2	El modelo GCSN	23
2.1	La necesidad de un mejor modelo	23
2.2	Introducción	23
2.3	Detalles	24
2.4	Diferencias con otros modelos	26
2.5	Fronteras especiales (singulares)	27
2.5.1	Definición de especialidad	28
2.6	Fronteras aleatorias (generales)	28
2.6.1	Frontera especial asociada	28
3	Lo que aporta esta tesis	31
II	Una solución: dinámica molecular	33
4	Introducción a la dinámica molecular	35
4.1	Simulación vs solución numérica	35

4.2	El papel de las simulaciones	35
4.2.1	¿Teoría o experimento?	36
4.3	Diferentes métodos para simular	37
4.4	Historia de la Dinámica Molecular	39
4.5	Dinámica Molecular actual	39
4.6	Limitaciones de la Dinámica Molecular	40
4.6.1	Sólo mecánica clásica	40
4.6.2	Realismo de los potenciales usados	41
4.6.3	Limitantes de tiempo	41
4.6.4	Limitantes de tamaño	41
4.6.5	Limitaciones en precisión al reproducir sistemas dinámicos	42
4.7	Monte Carlo	42
5	La maquinaria básica de la Dinámica Molecular	45
5.1	Configuración	45
5.1.1	Posiciones	45
5.1.2	Velocidades	51
5.2	El potencial	52
5.3	Ecuaciones de movimiento	53
5.3.1	2D ó 3D	55
5.3.2	Restricciones y constricciones	56
6	Ejecutando, midiendo y analizando	57
6.1	Usando unidades más simples	57
6.2	Integrando las ecuaciones de movimiento	58
6.2.1	Euler	59
6.2.2	Runge-Kutta	59
6.2.3	El algoritmo de Verlet	59
6.2.4	El algoritmo Salto de Rana	61
6.2.5	Algoritmo de Verlet Velocidad	61
6.2.6	Un poco más de Verlet	62
6.2.7	Los métodos Predictor-Corrector	62
6.2.8	Métodos que utilizan variables de varios instantes	63
6.2.9	Métodos que utilizan derivadas de alto orden	64
6.2.10	Comparación de los algoritmos de Verlet y de Predictor-Corrector	65
6.2.11	Otros métodos	65
6.2.12	Eligiendo un método	65
6.2.13	Eligiendo el tamaño de dt	66
6.3	Condiciones de frontera	67
6.3.1	Libres o de vacío	67
6.3.2	Para simular sistemas infinitos	68
6.3.3	Condiciones de frontera periódicas	69
6.3.4	De la simetría de la teselación	70
6.3.5	Fronteras de espejo	71
6.3.6	Condiciones de frontera esféricas	72
6.3.7	Del cálculo de las interacciones en sistemas con condiciones de frontera periódicas	72
6.3.8	Geometrías con superficies	74
6.3.9	Fronteras fijas y estocásticas	75
6.4	Hacia el equilibrio	75
6.5	Mirando los átomos	78
6.6	Midiendo el ensamble microcanónico	79
6.6.1	Energía potencial	81
6.6.2	Energía cinética	81
6.6.3	Energía total	81
6.6.4	Momento lineal	82
6.6.5	Momento angular	82
6.6.6	Función H de Boltzmann	82

6.6.7	Temperatura	83
6.6.8	Curva calórica	84
6.6.9	Densidad	84
6.6.10	Presión	85
6.6.11	Desplazamiento cuadrático medio	86
6.6.12	Energía libre	88
6.7	Midiendo la temperatura de fusión	88
6.8	Correlaciones en el espacio real	89
6.9	Correlaciones del espacio recíproco	91
6.10	Análisis dinámico	92
6.10.1	Función de auto-correlación espacio-tiempo	92
6.10.2	Función de correlación distintiva espacio-tiempo	94
6.10.3	Factor de estructura dinámico	95
6.11	Dinámica molecular como herramienta de optimización	96
6.11.1	Busqueda en línea	97
6.11.2	Máximo descenso	98
6.11.3	Gradiente conjugado	99
6.11.4	Métodos de Newton-Raphson	100
6.11.5	Recocido y templado	100
6.11.6	Metodología general de la minimización	101
6.12	Controlando la temperatura	102
6.12.1	Escalamiento directo	103
6.12.2	Baño térmico de Berendsen	105
6.12.3	Baño térmico estocástico	106
6.12.4	Método de retroalimentación de Nosé-Hoover	106
6.12.5	Otros métodos	108
6.13	Controlando la presión	108
6.13.1	Método de Berendsen	108
6.13.2	Método de Andersen de sistema extendido	109
6.13.3	Método de Parrinello-Rahman	109
6.13.4	Otros métodos	110
6.14	Simulando otros ensambles	110
6.14.1	Ensamble NVE (microcanónico)	110
6.14.2	Ensamble NVT (canónico)	111
6.14.3	Ensamble NPT (isotérmico-isobárico)	111
6.14.4	Ensamble NST	111
6.14.5	Ensamble NPH	111

7 Potenciales interatómicos 113

7.1	La aproximación de Born-Oppenheimer	114
7.2	Suponiendo una interacción más simple	115
7.3	Diseñando potenciales	115
7.4	Potenciales de pares	116
7.4.1	Tipo Coulomb	116
7.4.2	Potenciales Simples	116
7.4.3	Lennard-Jones	117
7.4.4	Morse	121
7.5	Los problemas con los potenciales de pares	121
7.6	Potenciales de muchos cuerpos para metales	122
7.6.1	Modelo del átomo embebido (EAM)	123
7.7	Potenciales de muchos cuerpos para semiconductores	124
7.7.1	Potencial de Stillinger-Weber	125
7.7.2	Potencial de Tersoff	126
7.8	Fuerzas de largo alcance y sumas de Ewald	126
7.9	¿Realmente son indispensables los potenciales?	127
7.10	Datos ab-initio por ajuste de fuerzas	127

8	Queremos un buen simulador	131
8.1	¿Que es bueno?	131
8.1.1	Bondad para la física	131
8.1.2	Bondad como programa	132
8.2	La parte más lenta	133
8.3	Calculando eficientemente las fuerzas, energías y presiones	134
8.4	Evitando la raíz cuadrada	134
8.5	Uso de tablas en lugar de funciones	135
8.5.1	Suavizando datos discretos	136
8.5.2	Derivadas numéricas	139
8.6	Potenciales desplazados	141
8.7	Eligiendo el radio de corte	141
8.8	Listas de vecinos	143
8.8.1	Lista de vecinos de Verlet	143
8.8.2	Subdivisión en celdas	145
8.8.3	Celdas mínimas	146
8.9	Método de muchos intervalos	146
8.10	Usando varios procesadores	147
8.11	Usando varias computadoras	147
9	Diseño y funciones del programa	149
9.1	Que esperamos del programa	149
9.1.1	Lo que si debe hacer	149
9.1.2	Lo que no debe hacer	150
9.2	¿Que maquina vamos a usar?	150
9.3	¿En que lenguaje?	150
9.4	Objetos	151
9.4.1	¿Como se crean?	152
9.4.2	Interacciones entre objetos	152
9.4.3	Ventajas de usar objetos	152
9.4.4	Desventajas de usar objetos	153
9.4.5	Objetos en nuestro programa	153
9.5	Los átomos	157
9.5.1	Selección y marcado	158
9.6	Flujo	159
9.6.1	Etapas	159
9.6.2	Diagramas de flujo	161
9.7	Manejo de unidades	165
9.7.1	Internas	165
9.7.2	Conversion de unidades internas \longleftrightarrow externas	165
9.8	Instrucciones, el archivo ini	166
9.8.1	Elementos del archivo ini	167
9.8.2	Ejemplos de instrucciones	174
9.9	Formatos de entrada y salida	176
9.9.1	Configuración	176
9.9.2	Películas	177
9.9.3	Potenciales	177
9.9.4	Mediciones	178
9.9.5	Bitácora (Log)	178

IV	Resultados	179
10	La ecuación de estado <i>universal</i> para metales	181
10.1	Evidencia experimental	184
10.2	Nos preguntamos	184
10.3	El procedimiento (algoritmo) para estudiarlo	185
10.4	Resultados	185
10.5	Conclusiones	188
11	Fronteras de grano	189
11.1	Los experimentos	189
11.1.1	Nuestras intenciones	189
11.1.2	Fronteras de interés	190
11.1.3	Método de enfriado	190
11.1.4	Método de temperatura constante	191
11.1.5	Potenciales	192
11.1.6	Condiciones de periodicidad	193
11.1.7	Método de integración	194
11.2	Resultados	194
11.2.1	Análisis de los resultados	200
11.2.2	Conclusiones de fronteras de grano	204
11.3	Conclusiones	206
11.4	A futuro	212
V	Apéndices	215
A	Hardware y software	217
A.1	La computadora	217
A.1.1	¿Que computadora?	217
A.1.2	Almacenamiento	218
A.1.3	El CPU	218
A.2	Lenguaje	219
A.2.1	Basic	221
A.2.2	¿Fortran o C?	221
A.2.3	C++	222
A.2.4	De Pascal y Java.	223
A.3	Software existente	224
B	Formatos	225
B.1	Configuraciones	225
B.1.1	Formato PDB	225
B.1.2	Formato XYZ	226
B.1.3	Para transformar de un formato a otro	226
B.2	Películas	226
C	Constantes y factores de conversión	227
C.1	Factores de conversión	227
C.2	Constantes	227
C.2.1	Parámetros para el potencial de Lennard-Jones de argón	227
D	Talachas Matemáticas	229
D.1	Función del virial de Clausius	229
D.2	Errores	230
D.2.1	Errores de integración	230
D.2.2	Error estadístico	231

Índice de Figuras

1-1	Fotografía de alta resolución de una frontera heterofase niobio-safiro.	4
1-2	Fronteras homofase y heterofase.	5
1-3	Dislocación de borde.	5
1-4	Frontera de bajo ángulo.	6
1-5	Imágen de una frontera de bajo ángulo.	7
1-6	Ejemplo de una frontera de gemelos.	8
1-7	Frontera escalonada.	9
1-8	Período de una frontera.	10
1-9	Frontera cercana a una CSL.	10
1-10	Camino escalonado no simétrico.	11
1-11	Planos coincidentes.	12
1-12	Energía potencial de burbujas.	13
1-13	Dislocación en burbujas.	13
1-14	Dislocaciones de fronteras entre burbujas.	14
1-15	Ejemplo de una O-lattice y la red de dislocaciones.	14
1-16	Definición de un punto-O.	15
1-17	Cálculos de la energía de una frontera como función de la desorientación.	17
1-18	Histograma de fronteras observadas como función de la desorientación.	17
1-19	Separación entre dislocaciones.	18
1-20	Variación de la anisotropía en la difusión como función de la desorientación.	19
1-21	Densidad de precipitados como función de la desorientación.	20
1-22	Fractura intergranular	21
1-23	Movilidad de una frontera de grano.	22
1-24	Energía de activación para la movilidad como función de la desorientación.	22
2-1	Diferencias en la construcción de fronteras.	24
2-2	Frontera aleatoria antes y después de una relajación estática.	25
2-3	Gráfica normalizada de la densidad atómica contra la desorientación.	25
2-4	Estructuras de fronteras de grano con $\Sigma 5$, $\Sigma 41$, $\Sigma 61$, $\Sigma 17$ y $\Sigma 37$	26
2-5	Fronteras de grano asociadas a $\Sigma 5$	27
4-1	Diagrama de la relación entre diferentes métodos de simulación.	38
5-1	Celda FCC con su celda básica.	46
5-2	Celda primitiva FCC.	46
5-3	Cloruro de Sodio.	47
5-4	Diamante.	47
5-5	Secuencia para formar cristales FCC o HCP.	48
5-6	Distribución de pares para 30.000 átomos en FCC.	49
6-1	Pasos sucesivos de diferentes métodos de Verlet.	61
6-2	Divergencia de trayectorias según diferentes condiciones iniciales.	67
6-3	Deriva en la energía para diferentes algoritmos y dt	68
6-4	Fluctuaciones de la energía para varios algoritmos de integración.	69
6-5	Condiciones de frontera periódicas y conservación de la masa.	70
6-6	Celdas periódicas no cúbicas.	71
6-7	Frontera de espejo.	72

6-8	Convención de imagen mínima.	73
6-9	Interacciones como función del radio de corte.	73
6-10	Convención de imagen explícita.	74
6-11	Fase de equilibrio.	76
6-12	Fluctuaciones en la energía para sistemas de diferentes tamaños.	78
6-13	Estabilidad de un sistema en equilibrio ante cambios de energía.	79
6-14	Trayectorias atómicas para sólido y líquido.	80
6-15	Diferencia en el desplazamiento cuadrático medio para sólidos y líquidos.	87
6-16	Frontera sólido-líquido para medir temperatura de fusión.	89
6-17	Función de distribución radial para diferentes estructuras.	90
6-18	Funciones de correlación de pares para dos estados diferentes.	90
6-19	Función de autocorrelación de espacio tiempo.	94
6-20	Función de correlación espacio-tiempo distintiva.	95
6-21	Curvas de nivel de una superficie de energía.	96
6-22	Minimización por búsqueda en línea.	97
6-23	Búsqueda en línea.	97
6-24	Minimización por máximo descenso y búsqueda en línea.	98
6-25	Minimización por máximo descenso sin búsqueda en línea.	99
6-26	Minimización por gradiente conjugado.	102
6-27	Minimización por enfriamiento.	103
6-28	Energías cinéticas controladas.	104
6-29	Energía potencial cuando la energía cinética se controla.	104
6-30	Escalones en la energía total al escalar velocidades.	105
6-31	Distribución de la energía cinética.	105
7-1	Forma de los potenciales de pares simples.	117
7-2	Coordenadas de dos osciladores.	118
7-3	Traslape de las distribuciones electrónicas al aproximarse dos átomos.	120
7-4	Efecto del principio de Pauli en la energía de dos átomos que se aproximan.	120
7-5	Potencial de Lennard-Jones y su fuerza derivada.	121
7-6	Potencial de átomo embebido.	125
8-1	Convergencia de energía cinética.	132
8-2	Método de Newton-Gregory de interpolación por diferencias adelantadas.	136
8-3	Efectos de cambiar el número de iteraciones de suavizado.	138
8-4	Efecto de cambiar el número de puntos para suavizar.	138
8-5	Problemas en la derivada numérica de potenciales discretos de archivo.	140
8-6	Cálculo numérico de segundas derivadas.	140
8-7	Lennard-Jones desplazado.	141
8-8	Energía potencial por ordenes de vecinos.	142
8-9	Diferentes métodos para calcular interacciones.	143
8-10	Lista de vecinos.	144
8-11	Separación y fuerza entre dos átomos.	145
8-12	Celdas mínimas.	146
8-13	Método de muchos intervalos.	147
9-1	Diagrama de relaciones de derivación de clases.	154
9-2	Celdas para recortar/marcar/ordenar los átomos.	155
9-3	Orden de los átomos dentro del arreglo <i>cluster</i>	158
9-4	Diagramas de flujo parte 1.	160
9-5	Diagramas de flujo parte 2.	162
9-6	Diagramas de flujo parte 3.	163
10-1	Energía de amarre para diferentes metales.	181
10-2	Energías de amarre para diferentes metales con energías escaladas.	182
10-3	Aproximación por un potencial armónico.	183
10-4	Evidencia experimental de una relación universal para la energía de amarre.	184
10-5	Energías de amarre de potenciales de pares y de átomo embebido.	186

10-6	Potencial triangular.	186
10-7	Potencial de Wyoming.	187
10-8	Energías de amarre de potenciales de pares con joroba y de otras figuras.	187
11-1	Diferentes parámetros de minimización de energía.	191
11-2	Diferentes potenciales para fronteras de grano.	192
11-3	Evolución de la energía total usando el potencial de Bristowe.	194
11-4	Evolución de la energía total usando el potencial de Rydberg.	195
11-5	Evolución de la energía total usando el potencial de YYYe-Cai.	195
11-6	Evolución de la temperatura usando el potencial de Bristowe.	196
11-7	Evolución de la temperatura usando el potencial de Rydberg.	196
11-8	Evolución de la temperatura usando el potencial de YYYe-Cai.	197
11-9	Cambios en la temperatura usando el potencial de Bristowe.	197
11-10	Cambios en la temperatura usando el potencial de Rydberg.	198
11-11	Cambios en la temperatura usando el potencial de YYYe-Cai.	198
11-12	Configuraciones de la frontera $\Sigma 13t2$ usando el potencial de Bristowe.	199
11-13	Configuraciones de la frontera $\Sigma 13t2$ usando el potencial de Rydberg.	200
11-14	Configuraciones de la frontera $\Sigma 13tcsl$ usando el potencial de Bristowe.	201
11-15	Configuraciones de la frontera $\Sigma 13tcsl$ usando el potencial de Rydberg.	202
11-16	Configuraciones de la frontera $\Sigma 5t2$ usando el potencial de Bristowe.	203
11-17	Configuraciones de la frontera $\Sigma 5t2$ usando el potencial de Rydberg.	204
11-18	Configuraciones de la frontera $\Sigma 13tcsl$ usando el potencial de Bristowe.	205
11-19	Configuraciones de la frontera $\Sigma 13tcsl$ usando el potencial de Rydberg.	206
11-20	Configuraciones de la frontera $\Sigma 13t2$ usando el potencial de YYYe-Cai.	207
11-21	Distribuciones de pares de la configuración $\Sigma 5t2$ inicial y final.	207
11-22	Diferencias entre las distribuciones de pares con o sin frontera $\Sigma 5t2$	208
11-23	Evolución de la energía total a diferentes temperaturas constantes de $\Sigma 13t2$	208
11-24	Evolución de la temperatura para simulaciones a temperatura constante.	209
11-25	Configuraciones finales de simulación a temperatura constante.	210
11-26	Configuraciones intermedias durante una simulación con energía constante.	211
11-27	Detalle de la simetría extra observada simulando NVE.	212
A-1	Diferentes niveles de programación.	220
D-1	Errores local y global al integrar.	230

Índice de Tablas

5.1	Ejemplos de combinaciones de coordenadas de vecinos en una configuración FCC	48
6.1	Unidades reducidas.	58
6.2	Coefficientes correctores de Gear (All87)	63
7.1	Comparación entre el potencial de Lennard Jones y resultados experimentales.	121
8.1	Efectos del radio de corte y corrección en la energía potencial	142
9.1	Conversion de unidades.	166
10.1	Parámetros de escala para algunos metales [Ros84].	183
A.1	Comparación entre diferentes lenguajes y sus versiones más populares (Erc97).	222
A.2	Programas útiles al programar simulaciones de dinámica molecular.	224
A.3	Sitios útiles al programar simulaciones de dinámica molecular.	224

0.1 Prólogo

0.1.1 Propósitos

El principal objetivo de esta tesis es tratar el problema de la estabilidad de las fronteras de grano generadas según el modelo GCSN. Sin embargo hay otros dos objetivos también importantes. 1) Proveer al grupo de estudio de propiedades de la materia de un programa versátil para simular por dinámica molecular. 2) Generar un manual de simulación por dinámica molecular y de fronteras de grano que permita a estudiantes de los últimos semestres de la carrera de física conocer los elementos clave necesarios para trabajar con estos temas.

Mi intención no es repetir lo que mencionan los libros en la materia, sino resaltar algunos de los detalles que los libros "pasan por alto", así como hacer algunas recomendaciones para su implementación en lenguaje C++ y al estudio de fronteras de grano. Pretendo reducir el tiempo requerido para que quienes realizarán su tesis de licenciatura simulando por dinámica molecular o trabajando con fronteras de grano participen activamente, logrando un mejor aprovechamiento de los futuros estudiantes.

Con esta tesis se generó un programa en C++ capaz de incorporar la mayoría de los algoritmos discutidos en los capítulos teóricos. Las posibilidades del programa son mucho mayores a las que se explotaron en la sección de cálculos de fronteras de grano, y mi intención es que algún estudiante futuro pueda utilizar esta tesis para comprender la forma en que fue creado el programa, para modificarlo, adaptarlo y extenderlo para hacer más simulaciones y entender mejor las fronteras de grano y el modelo GCSN.

A continuación describo el contenido de cada parte/capítulo y hago recomendaciones respecto a lo que se debe leer dependiendo de los intereses del lector.

0.1.2 Una visión panorámica

- Parte I - Presentación del problema.
 - Capítulo 1 - El problema de las fronteras de grano, los modelos que las describen, sus propiedades y las dificultades del campo.
 - Capítulo 2 - El model GCSN, lo que puede solucionar y la necesidad de la simulación por dinámica molecular como herramienta para estudiarlo.
 - Capítulo 3 - Los objetivos de esta tesis.
- Parte II- La dinámica molecular.
 - Capítulo 4 - Ideas generales de simulación, tipos, historia, virtudes y defectos.
 - Capítulo 5 - Maquinaria básica de dinámica molecular.
 - Capítulo 6 - Detalles de cada una de las partes, opciones, métodos, algoritmos y cálculos.
 - Capítulo 7 - Potenciales interatómicos. Razones de ser y distintas opciones.
- Parte III - El programa resultado/herramienta de esta tesis.
 - Capítulo 8 - Consideraciones y métodos para hacer buenas simulaciones.
 - Capítulo 9 - Diseño del programa, sus partes, la manera de usarlo y los detalles de su funcionamiento.
- Parte IV - Resultados.
 - Capítulo 10 - La ecuación de estado *universal* para metales. Estudios sobre su universalidad y aplicabilidad a simulaciones de fronteras de grano.
 - Capítulo 11 - Simulaciones de fronteras de grano. Simulaciones de enfriamiento para minimizar energía. Cálculo de temperatura de recristalización.

- Parte V - Conclusiones.
 - Capítulo 12 - Conclusiones.
 - Capítulo 13 - El futuro, consideraciones y propuestas.
- Parte VI - Apéndices.
 - Apéndice A - Discusión sobre los diferentes lenguajes usados para programar simulaciones. Las opciones en hardware y software para simular.
 - Apéndice B - Formatos estándar de archivo.
 - Apéndice C - Constantes y factores de conversión.
 - Apéndice D - Formalismos matemáticos.

¿Que debo leer?

Definitivamente este texto no es corto, para evitar que sea abandonado en favor de textos más pequeños, me permito recomendar las secciones que pueden ser de interés dependiendo de los propósitos del lector. Si la intención del lector es iniciarse en el estudio de fronteras de grano basta con la primera parte y las conclusiones. Para aquellos con la intención de usar esta tesis como manual para aprender dinámica molecular recomiendo la segunda parte y el capítulo 8 de la parte III. Para los que buscan los temas relacionados con programación y aplicación de métodos orientados a objetos en simulación el capítulo 9 de la parte III y el apéndice A son los indicados. Los interesados en los potenciales de interacción y sus aplicaciones encontrarán adecuado el capítulo 10. Finalmente para quienes sean víctimas de los mismos intereses que yo y quieran estudiar la simulación por dinámica molecular aplicada a fronteras de grano sugiero leer el texto completo.

Parte I

El problema

1. FRONTERAS DE GRANO

Los materiales que están compuestos de agregados de cristales se denominan *policristalinos*. Cada cristal se llama *grano*, mientras que las regiones de encuentro entre diferentes granos se denominan *fronteras de grano*.

1.1 Importancia científica y tecnológica

“*Algunos días se es el perro, otros el poste.*” Anónimo.

La motivación para estudiar fronteras de grano proviene, como en muchas ocasiones, de dos fuentes:

- Una necesidad tecnológica de aprovechar las propiedades de los policristales.
- Una búsqueda de conocer y entender los materiales de que está hecho el mundo.

La mayoría de los materiales que utilizamos se encuentran en formas policristalinas. No solamente por las enormes dificultades para crear un monocristal, sino porque sus propiedades son diferentes y pueden ser más útiles. Algunos ejemplos son partes de aviones, de automóviles, de reactores nucleares y de piezas electrónicas. Muchas veces las fronteras de grano son regiones de particular debilidad en el material, y fallas catastróficas como fracturas ocurren más probablemente a lo largo de ellas (sección 1.4). El costo para la sociedad de tales fallas es enorme, no solo en miles de millones de dólares, sino en vidas.

Ejemplos de estas situaciones son los cuerpos de aviones, piezas de reactores nucleares, motores de auto y avión y electrónica. A pequeña escala un ejemplo es la falla de circuitos de computadora, que se atribuye a las fronteras de grano en las conexiones eléctricas. A mayor escala, las paletas del compresor en las turbinas de un avión son reemplazadas periódicamente según mediciones estadísticas de su durabilidad, y no en el momento en que pueden presentar fallas. El resultado es cambiar paletas que aún tienen una larga vida útil por delante a cambio de minimizar la posibilidad de falla durante el vuelo.

¿Porque no se aprovechan los modelos de fronteras de grano para determinar el momento adecuado para reemplazar las costosas paletas y evitar las fallas en las conexiones eléctricas?. La respuesta es simple: porque no se tienen modelos adecuados de las fronteras. No se saben las implicaciones de las propiedades de las fronteras en las del material completo. En consecuencia no se pueden controlar adecuadamente ni se conocen métodos de medición que puedan determinar las causas de falla.

La teoría se encuentra tan estancada que no se tienen definiciones precisas para las clasificaciones empíricas de las fronteras de grano.

Los ingenieros han tratado de darle la vuelta al problema invirtiendo cientos de millones de dólares en experimentos, acumulando una enorme cantidad de datos empíricos. La mesa está puesta para los científicos. Un problema interesante, importante, con implicaciones económicas evidentes y grandes cantidades de evidencia experimental.

Uno de los primeros objetivos es describir la estructura atómica de las fronteras de grano, para ello se han desarrollado diferentes modelos. En la sección 1.3 describo los más importantes, mientras que en la sección 1.4 presento una síntesis de las propiedades de las fronteras de grano.

1.2 Clasificación

Hay diferentes clasificaciones para las fronteras [Fin92]:

1. Según los materiales que dan lugar a la frontera (figura 1-2).
 - (a) Fronteras de la misma fase (Homofase) - Interfaces entre granos de idéntica estructura cristalina y composición.



Figura 1-1 Imagen de alta resolución invertida de una interfase heterofase Nb/Al_2O_3 (niobio-safiro). Los planos de las redes son claramente identificables en ambos materiales. Las regiones de buen encaje (M), mal encaje (D) y los escalones (S) están marcados para facilitar la identificación [Fin92, pág 541].

(b) Fronteras de diferentes fases (Heterofase) - Interfaces entre regiones de diferente estructura cristalina o composición. Usualmente se conocen como *fronteras de fase*.

2. Según sus propiedades.

(a) Fronteras especiales - Aquellas con propiedades particulares. No hay un consenso aún de su definición, aunque hay acuerdo en que una Σ pequeña es un buen criterio de especialidad ($\Sigma = \frac{\text{volumen de la celda unidad CSL}}{\text{volumen de la celda unidad del cristal}}$. Ver sección 1.3.5).

(b) Fronteras aleatorias - Todas las fronteras que no son especiales.

3. Según la magnitud de la rotación.

(a) Fronteras de bajo ángulo - Para diferencias de orientación de menos de $10-15^\circ$ (dependiendo del caso específico).

(b) Fronteras de alto ángulo - Para diferencias de orientación de más de $10-15^\circ$ (dependiendo del caso específico).

4. Según la dirección del vector sobre el que se hizo la rotación.

(a) Fronteras de rotación (Twist) - Cuando el eje de rotación es perpendicular al plano de la frontera.

(b) Fronteras de inclinación (Tilt) - Cuando el eje de rotación es paralelo al plano de la frontera.

1.3 Modelos

“Una teoría solo puede ser correcta o incorrecta. Un modelo tiene una tercera posibilidad: puede ser irrelevante” Manfred Eigen [Bri99].

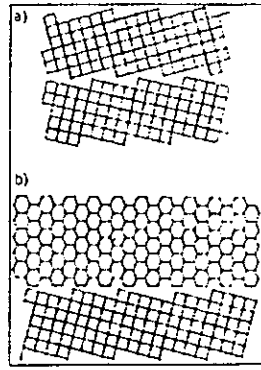


Figura 1-2 Diferentes fronteras según los cristales que la forman. a) Frontera homofase. b) Frontera Heterofase [Pum76].

1.3.1 Fronteras de bajo ángulo:

Una manera simple, exitosa y generalmente aceptada para entender las fronteras de grano de bajos ángulos sugerida originalmente por Bragg [Bra40] y Burgers [Bur40] y desarrollada por Read y Shockley [Rea50] explica las variaciones entre la configuración cristalina y aquella de la frontera por medio de la inclusión de dislocaciones. En la figura 1-3 se muestra una dislocación de borde, y la forma en que modifica la estructura local (y global) de un cristal, pudiendo dar lugar a una rotación de una parte del cristal respecto de la otra. Un ejemplo de la forma en que una frontera de bajo ángulo es explicada como resultado de dislocaciones se muestra en la figura 1-4. La separación entre las dislocaciones que dan lugar a un ángulo de rotación θ es

$$D = \frac{b}{2 \sin(\frac{\theta}{2})} \approx \frac{b}{\theta} \quad (1.1)$$

donde b es el vector de burgers de las dislocaciones. Los experimentos apoyan este modelo, por ejemplo, la figura 1-5 muestra la distribución de dislocaciones a lo largo de una frontera de bajo ángulo observada con un microscopio electrónico en donde hay coincidencia entre las mediciones directas y las predicciones de la ecuación 1.1.

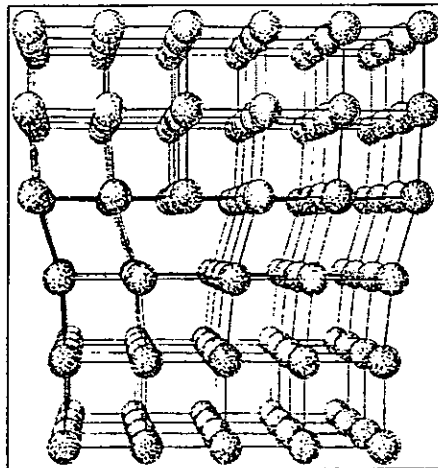


Figura 1-3 Estructura de una dislocación de borde. La deformación puede pensarse como causada por insertar un plano de átomos extra en la mitad superior. Átomos en esa mitad necesitan comprimirse por la inserción, aumentando la energía con respecto a la mitad inferior. Notese cómo los átomos más cercanos al defecto tienen una coordinación (número de vecinos) muy diferente a la cristalina. [Kit86]

Para una descripción de la naturaleza y propiedades de las dislocaciones en cristales sugiero el libro de Kittel [Kit86].

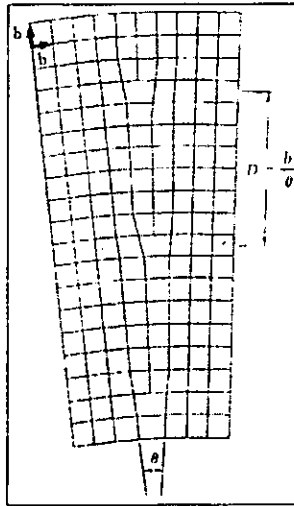


Figura 1-4 Frontera de bajo ángulo, según Burgers. [Kit86]

1.3.2 Fronteras de alto ángulo:

Dado que la distancia entre las dislocaciones es una función decreciente del ángulo, conforme este aumenta la densidad de dislocaciones será mayor. Cuando el ángulo ha llegado a unos $10 - 15^\circ$ la distancia en las dislocaciones es tan pequeña (y alta su densidad) que pierden su identidad como defectos, lo que tiene consecuencias notables en las propiedades de la frontera.

Los modelos de dislocaciones explican satisfactoriamente las fronteras de bajo ángulo y algunos aspectos de las de alto ángulo, pero no logran predecir la existencia de fronteras de alto ángulo especiales, por lo que son necesarios modelos más generales. Hasta hoy se han desarrollado varios, con diferentes niveles de complejidad y éxito. A continuación doy una breve explicación de los principales:

1.3.3 Modelo de estructura amorfa:

Originalmente se pensaba en las fronteras como capas de material amorfo [Bri98] [Qui05] [Ros13a] [Ros13b]. Sin embargo, si así fuera, todas las fronteras deberían tener las mismas propiedades. Hay evidencia experimental de que las propiedades de las fronteras cambian tanto con el ángulo de desorientación (diferencia de orientación entre los cristales que dan lugar a las fronteras) como con el plano de la frontera [Gle72] y es recomendable abandonar esa idea (ver sección 1.4.2).

Si una frontera de grano fuese amorfa tendría la misma energía libre que el cristal tiene más allá del punto de fusión, en contradicción con la evidencia experimental [McI65]. Esta propiedad es resultado de suponer los cristales como teselaciones infinitas del espacio, ya que en sistemas finitos (por ejemplo clusters) se ha mostrado la posibilidad de tener estructuras amorfas de menor energía que las de mayor simetría [Gar98].

A pesar de esto el modelo de estructura amorfa no es una tontería, predice cualitativamente algunas propiedades, como el deslizamiento observado a altas temperaturas y la fragilidad a bajas.

1.3.4 Modelo de Islas

Consiste en la suposición de que una frontera de grano está compuesta de dos tipos diferentes de regiones, uno donde la estructura es prácticamente cristalina (y por lo tanto con baja energía) y otro en donde es amorfa. Las regiones aproximadamente cristalinas serán las regiones de muy buena coincidencia entre ambos granos de cristal y seguramente serán pequeñas, rodeadas del "mar" amorfo.

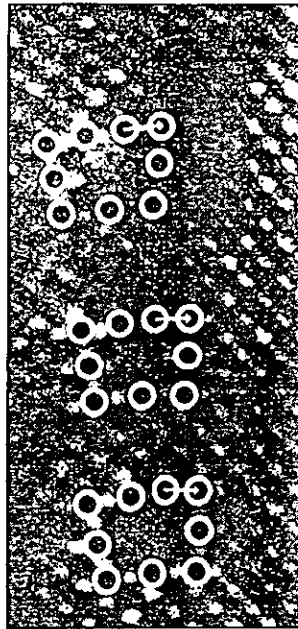


Figura 1-5 Imagen de microscopio de electrones de una frontera de bajo ángulo en molibdeno. [Kit86]

El modelo fué propuesto por Mott [Mot48] y ligeramente modificado por Ke [Ke49] para explicar resultados experimentales que mostraban altos valores de la entropía durante la migración de una frontera de grano. Según el modelo la migración ocurre cuando algunos átomos de la orilla de la región de buen arreglo se desordenan, mientras que del otro lado de la isla algunos átomos del mar amorfo se acomodan uniéndose a la isla. El resultado neto es el desplazamiento de la isla. Gifkins [Gif67] redujo los mares amorfos del modelo hasta llevarlos a canales entre las islas cristalinas.

1.3.5 Modelo de sitios de coincidencia (CSL)

Dos redes cristalinas A y B que se encuentran en una frontera son extendidas en forma imaginaria de manera que ambas cubran todo el espacio. Una de las redes, digamos A, es trasladada de manera que un punto coincida con uno de la red B. A este punto lo llamamos O por considerarlo desde ahora como el origen de coordenadas. Es posible que no haya más puntos de ambas redes que coincidan, pero en 1926 Friedel [Fri26] mostró que para algunas orientaciones especiales sí habrá mas sitios de coincidencia. Los sitios de coincidencia (de haber algunos) forman una red conocida como la Red de Sitios de Coincidencia (CSL por las siglas en inglés de *Coincidence Site Lattice*). Por supuesto, la CSL tiene un período mayor que los cristales que le dan lugar [Fin92].

Se observa experimentalmente que las fronteras de grano ocurren más en orientaciones que pueden ser asociadas con una CSL. Un ejemplo simple es el de la inclinación simétrica en un cristal cúbico, del cual un caso especial es el llamado gemelo (111) en un cristal FCC. El acomodo ABCABC... en el gemelo se invierte al pasar por la frontera. En un gemelo real hay una pequeña traslación de las dos partes porque el espacio entre los planos (111) cerca de la frontera no mantiene su valor exactamente. Después de quitar la traslación normal a la frontera, si los cristales son imaginados como interpenetrados, cada tercer plano (111) en un cristal coincide con un plano (111) en el otro cristal. La CSL en este caso tiene puntos cada tercer plano (111) (ver sección 5.1.1). La figura 1-6 muestra un ejemplo de una frontera de gemelos.

Con este modelo se puede clasificar a las fronteras de grano según la razón Σ del volumen de una celda unidad de la CSL al volumen de la celda unidad del cristal A o B:

$$\Sigma = \frac{\text{volumen de la celda unidad CSL}}{\text{volumen de la celda unidad del cristal}} \sim \frac{1}{\text{Fracción de coincidencias entre ambas redes}} \quad (1.2)$$

Notese que en la ecuación 1.2 el inverso de la fracción de coincidencias entre ambas redes es mayor o igual que uno (la igualdad ocurriendo solamente en el caso de la frontera trivial). Un valor bajo para Σ implica un gran número de

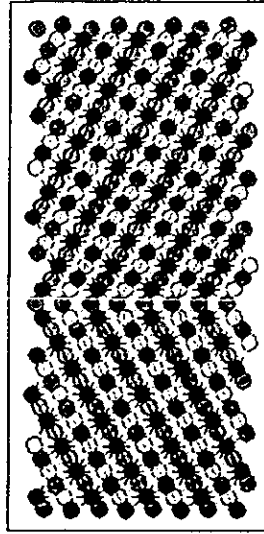


Figura 1.6 Ejemplo de una frontera de gemelos (twin) [Ame99]

coincidencias entre las redes interpenetradas; $\Sigma = \infty$ significa una orientación completamente inconmensurada o al azar. Fronteras con Σ relativamente pequeña son llamadas "fronteras de coincidencia" y por lo general son asociadas con propiedades físicas especiales como baja energía y alta movilidad.

En un monocristal podemos definir la frontera trivial cuando la desorientación entre los granos es igual a 0° , la CSL es idéntica a ambas redes y $\Sigma = 1$. Fuera de ese ángulo las máximas coincidencias en un sistema cúbico ocurren para $\Sigma = 3$, en donde un tercio de los átomos de un cristal coinciden con los del otro.

Según Ranganathan [Ran66] una CSL es una red tridimensional generada por una rotación de ángulo θ sobre el eje $[hkl]$ sí:

$$\theta = 2 \tan^{-1} \left(\frac{y}{x} \right) \sqrt{N} \quad (1.3)$$

donde x e y son enteros y

$$N = h^2 + k^2 + l^2. \quad (1.4)$$

De aquí se puede calcular

$$\Sigma = x^2 + Ny^2 \quad (1.5)$$

pero por convención si el resultado es un número par se divide entre dos hasta obtener un número impar.

Por el gran número de simetrías de los sistemas cúbicos una rotación puede ser descrita por 24 diferentes parejas (eje, ángulo), por ejemplo, una rotación de 53.1° sobre un eje $[001]$ también puede ser descrita como una rotación de 180° sobre el eje $[\bar{1}20]$. Una frontera de gemelos comúnmente descrita como una rotación de 180° sobre $[111]$ o $[11\bar{2}]$ es generalmente descrita por microscopistas como una rotación de 70.5° sobre $[1\bar{1}0]$. Es costumbre identificar las fronteras ya sea por el valor de Σ asociada con la densidad de puntos en común entre ambas redes o por el ángulo de rotación sobre el eje con los más bajos índices de Miller.

A un plano de la frontera que es también un plano de la CSL se le puede asociar además de una Σ , una cierta densidad $D \sim \Sigma^{-1}$ de sitios de coincidencia por unidad de área, que puede ser relevante para las propiedades estructurales [Pon74]. Originalmente se pensaba que fronteras con un valor grande para D tendrían una baja energía libre por tener un gran número de átomos en posiciones de coincidencia; hoy en día se sabe que esto no necesariamente es una regla general pues en fronteras generales las coincidencias geométricas se ven destruidas por traslaciones relativas de ambos granos o por acomodos atómicos (relajación) en la interface. Sutton y Baluffi en 1987 [Sut87] llegaron a la conclusión de que valores pequeños para Σ o algún otro criterio puramente geométrico no parecía poder estar correlacionado directamente con propiedades físicas medidas. Este es un punto sumamente importante al que hemos de volver repetidamente.

Algunos autores [Kro49] [Ran65] [Bra64] han sugerido que si las fronteras de CSL deben ser de alta coincidencia D (baja Σ) para tener una baja energía, entonces que la interface siga un plano con un alto número de coincidencias debe ser favorecido, y en ocasiones esto puede llevar incluso a tener una frontera escalonada (figura 1-7).

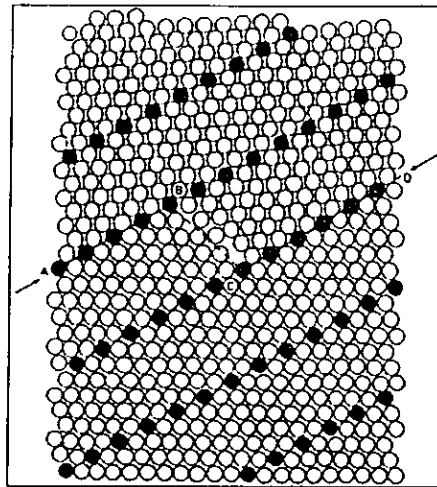


Figura 1-7 Modelo de una frontera $\Sigma = 11$ CSL en un bicristal bidimensional bcc. La frontera sigue el camino A, B, C, D para maximizar el camino a lo largo de altas densidades de puntos de coincidencia. Los puntos negros son sitios de coincidencia. [Pum76]

Unidades estructurales

La propiedad esencial de las fronteras CSL según Bishop y Chalmers [Bis68] es la presencia de pequeños dominios atómicos periódicos cuyo tamaño es característico para cada CSL y que no necesariamente es constante en una frontera.

Los modelos que mencioné hasta ahora están basados en consideraciones geométricas, si ahora pasamos a consideraciones termodinámicas de mínima energía, podríamos pensar en que una frontera estuviera gobernada por el principio de que la energía libre de un sistema debe ser mínima. Bajo esta luz tenemos las opciones:

- Los átomos de la frontera pueden moverse libremente minimizando su energía libre.
- Mover ligeramente ambos cristales acercándolos, aumentando la cantidad de átomos que forman parte de la frontera, disminuyendo el volumen del sistema y resolviendo el problema de la baja densidad en la frontera.

En esta variación del modelo de sitios de coincidencia además del énfasis en el enfoque termodinámico, se superan algunas de las objeciones a las fronteras CSL. Por ejemplo, en la figura 1-8 se observa como átomos que casi coinciden (átomos F y G), localizados junto a átomos de coincidencia (átomos A), están en estados de muy alta energía por encontrarse en una región sobrepoblada. El modelo propone reducir la energía libre eliminando el átomo G. *Todos los átomos pertenecen a uno u otro cristal [Lob76].*

También es posible reducir la energía sin necesidad de eliminar los átomos G, aplicando una traslación rígida a uno de los granos. Una propiedad geométrica que se preserva, aunque se realicen estas transformaciones y se eliminen las coincidencias exactas de átomos, es la periodicidad de la frontera. Aparentemente existen algunas confirmaciones experimentales de este tipo de traslaciones en fronteras con $\Sigma = 3$ [Pon74].

Simulaciones de [Wei69] [Wei70] donde relajan fronteras han confirmado que no necesariamente debe haber átomos en el plano de la frontera y han mostrado que la estructura debe consistir de unidades atómicas que se repiten con el mismo período que los bordes [Bis68].

A pesar de que la periodicidad en el plano de la frontera es la misma que en el caso CSL, la relajación de la frontera cambia los espaciamientos. Por lo tanto la CSL tridimensional construida con las distancias originales de la red ya no tiene sentido en esta región relajada. El tamaño de estas unidades estructurales repetidas se espera que crezca cuando el plano de la frontera se desvía de la posición simétrica [Cha71].

Pequeñas desviaciones de las desorientaciones CSL

El modelo de CSL sería aplicable a muy pocos casos si las estructuras especiales ocurrieran sólo cuando los cristales tienen la orientación relativa exacta correspondiente a una CSL. Algunos autores afirman que hay evidencia de que las propiedades de las fronteras CSL pueden extenderse a desviaciones cercanas [Gle72]. La explicación de este fenómeno es igual a la de las fronteras de bajo ángulo. Se supone una malla de dislocaciones con vectores de burgers

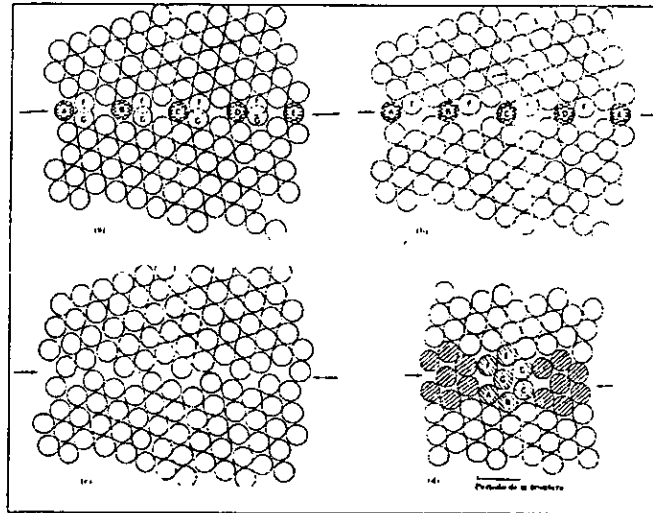


Figura 1-8 Modelo bidimensional de una frontera $\Sigma = 7$ CSL en un cristal HCP. [Pum76]

que llevan los átomos a los sitios de coincidencia. El efecto de esta malla es, nuevamente, condensar los malos encajes de coincidencia en líneas que separan "islas" de buen encaje.

En términos del modelo de unidades estructurales, agregar una red de dislocaciones es equivalente a agregar, periódicamente, una unidad característica de otra CSL. Bishop y Chalmers en 1968 dan un ejemplo de esto (figura 1-9). De hecho, ellos van más allá, postulando que la combinación adecuada de estas unidades puede generar cualquier frontera de alto ángulo. Hay simulaciones que apoyan esta idea al mostrar que todas las fronteras (descritas según este modelo) están relacionadas con, al menos, una frontera del conjunto al que llamamos "especial" (ver sección 1.4.2) y que entonces no existen las fronteras completamente azarosas. Solamente con el uso de dislocaciones y sin el modelo de unidades estructurales no es posible alejarse (en ángulo) mucho de las fronteras de muy buena coincidencia hacia las fronteras de alto ángulo, por la misma razón que antes era imposible explicar las fronteras de alto ángulo, porque al aumentar la densidad de dislocaciones estas se acercan, llegando al punto de perder su identidad. No está claramente definido a qué ángulo dejan de tener sentido las dislocaciones, ni el valor máximo de Σ para el que una CSL tiene sentido, pero no cubren los 180° de posibles desorientaciones.

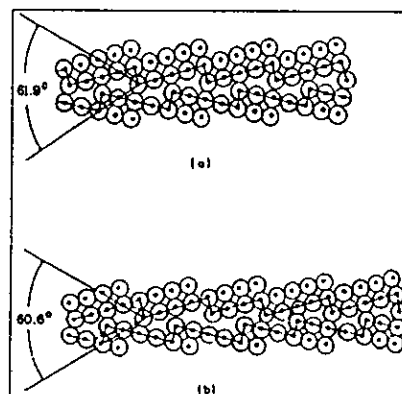


Figura 1-9 a) Frontera de $61.9^\circ/[100]$ $\Sigma = 17$ de cristal FCC convertida en b) frontera $60.6^\circ/[100]$ por la introducción periódica de una unidad diferente. [Pum76]

Fronteras con baja Σ pueden tener dificultades particulares [Wei72], un ejemplo es la figura 1-10 para $\Sigma = 39$ en donde no necesariamente la frontera simétrica DE es la de menor energía, pues también existe la posibilidad de que la frontera escalonada DAAE sea de menor energía. Para otras fronteras puede presentarse el mismo problema, en general difícil de tratar.

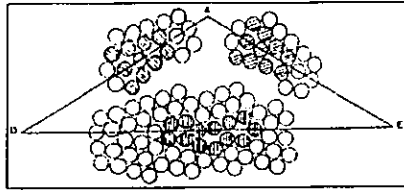


Figura 1-10 Estructuras de mínima energía calculadas ($\Sigma = 39$). Para el camino simétrico D-E el período es largo y la energía alta. Para la alternativa D-A-E el período es más corto y la energía puede ser menor. Los átomos sombreados son la unidad repetida. [Pum76]

1.3.6 Modelo de Casi Coincidencia

Bishop y Chalmers en 1967 [Bis68] hicieron notar que no necesariamente sólo pueden existir las diferencias de orientación que coinciden con una frontera CSL, sino que basta con que haya planos commensurados, y a diferencia de las fronteras CSL la coincidencia ocurra sólo para planos específicos. Es a esta situación a la que se le da el nombre de Frontera de Casi Coincidencia (NC por las siglas en inglés de *Near Coincidence*).

Es muy conocido el caso en que la no coincidencia entre dos redes es lineal, en donde puede ser absorbida por una red de dislocaciones.

Johannesson y Thölen hicieron notar en 1970 [Joh70] que, al menos para el caso de las fronteras de inclinación entre estructuras cúbicas [111], los períodos de casi coincidencia son muy cortos, y que esto probablemente se extiende a otras fronteras entre sistemas cúbicos; lo que podría significar que hay tantas fronteras NC que no hay necesidad de agregar una red de dislocaciones para mantenerlas al variar el ángulo. Proponen también que las fronteras NC probablemente sean más importantes en fronteras entre cristales no cúbicos o entre cristales de diferentes estructuras (fronteras de grano de heterofase) en los que hay menos redes CSL tridimensionales posibles.

1.3.7 Modelo de Planos Coincidentes

Diferentes autores (entre ellos [Pum72]) proponen que una frontera de alto ángulo a través de la cual un sólo conjunto de planos de bajo índice es continuo es de baja energía. Esto puede ocurrir en cualquier cristal, siempre y cuando los planos estén suficientemente cerca. En cristales FCC los planos posibles son: {200}, {220} y {110}.

La diferencia principal entre el modelo CSL y el PM (por las siglas en inglés de *Plane Matching*) es que mientras el modelo PM requiere que todas las fronteras creadas por rotaciones sobre un eje de bajo índice sean de baja energía, el modelo CSL requiere que sólo aquellas con desorientaciones específicas sean de baja energía. Una frontera PM puede ser considerada formalmente como un caso límite de una frontera CSL cuando $\Sigma \rightarrow \infty$ [War74].

La figura 1-11 muestra cómo ocurren estas coincidencias de planos para fronteras de a) inclinación y b) de rotación. El modelo propone que la estructura de la frontera en la dirección paralela a las líneas es similar a la de una frontera al azar de alto ángulo. En términos de una CSL, las líneas de mal ajuste son dislocaciones con vector de Burgers igual al espaciamiento interplanar.

El modelo fué propuesto por Pumphrey para explicar experimentos en que se observaban defectos estructurales aparentemente fuera de lo aceptable según la teoría de CSL. Pumphrey también propone que la coincidencia de planos es provocada por interacciones atómicas en la frontera que van más allá de primeros vecinos. Otra de las ventajas de este modelo (figura 1-11) es que el agregar defectos a la estructura de la frontera permite eliminar las tensiones a larga distancia que no pueden ser evitadas por dislocaciones porque si los planos tienen altos índices de Miller estas tendrían que encontrarse muy cercanas entre sí y no deberían de ser visibles en fotografías por microscopía electrónica o de campos de iones, contradiciendo lo observado experimentalmente.

Balluffi y Schober [Bal72] critican esta teoría diciendo que no es más que una manera diferente de describir fronteras por coincidencia, además de que una frontera PM hace predicciones ambiguas de los vectores de Burgers. Según ellos es posible distinguir entre ambas teorías experimentalmente en fronteras cercanas a la coincidencia, pues mientras PM predice fronteras rectas CSL predice fronteras escalonadas.

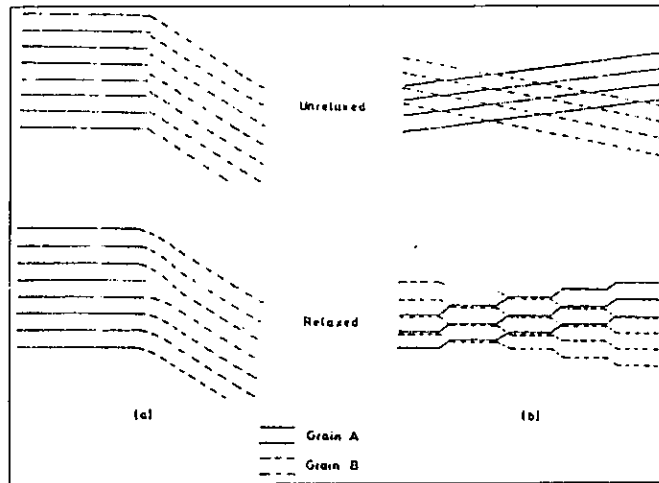


Figura 1-11 La forma en que se acomodan las desorientaciones de un conjunto de planos de bajo índice para fronteras de a) inclinación y b) rotación. [Pum76]

1.3.8 Modelo de burbujas.

La base del modelo es la siguiente observación de Ishida "Burbujas de jabón del mismo tamaño flotando en la superficie de agua forman un arreglo bidimensional con propiedades muy parecidas a las de metales" [Ish76].

Este modelo es un tanto peculiar, sobre todo porque no es exactamente un modelo de frontera de grano, sino una forma de visualizar defectos cristalinos y estudiar algunas de sus propiedades. Lo incluyo en esta recopilación porque ha sido usado para estudiar fronteras de grano de bajo ángulo y porque me parece fino, elegante y bonito.

El primero en examinar el modelo fué Bragg en 1942 [Bra42]. Lomer y Nye [Lom52] lo utilizaron para mostrar el arreglo de dislocaciones que dan lugar a una frontera de bajo ángulo, y que el modelo de islas puede explicar algunas fronteras de alto ángulo. Se estudiaron a profundidad los alcances del modelo, cuando Bragg y Nye [Bra47] desarrollaron una teoría para extender el arreglo a tres dimensiones. Fukushima y Ookawa [Fuk57] estudiaron la influencia de vibraciones térmicas en la estructura de la frontera aplicando vibraciones mecánicas a las burbujas.

Curiosamente al principio el modelo fué utilizado principalmente para ilustrar libros porque no existían las técnicas experimentales que pudieran corroborar sus predicciones de la estructura de la frontera [Ish76].

Con el invento del microscopio de campo de iones (FIM por las siglas del inglés *Field-Ion Microscope*) pudieron estudiarse las estructuras de la frontera con una precisión de hasta 3 Å [Bra64], con el éxito de mostrar algunas coincidencias entre los experimentos y las predicciones del modelo de burbujas. Más tarde con el advenimiento de las simulaciones por computadora se extendieron las coincidencias, colocando al modelo de burbujas en un sitio intermedio entre los experimentos con microscopio de campo de iones con resolución de algunos átomos, y las simulaciones por computadora, con excelente precisión posicional pero grandes aproximaciones.

El trabajo con burbujas tiene ventajas importantes: es muy visual, fácil de hacer, fácil de entender y muy económico. Sin embargo, tiene debilidades importantes:

1. Sólo es bidimensional - Se han desarrollado modelos tridimensionales pero no son muy prácticos, por lo que sólo pueden usarse para estudiar fronteras de inclinación [Fuk57].
2. No permiten estudiar efectos térmicos fácilmente - El modelo estacionario de burbujas corresponde a una temperatura cristalina de 0°K, útil en muchas circunstancias, pero no en todas. Aunque las vibraciones mecánicas se parecen a vibraciones térmicas [Fuk55], las frecuencias atómicas ($\sim 10^{12}$ s) son inalcanzables.
3. La naturaleza de las interacciones - La fuerza atractiva entre burbujas proviene de la tensión superficial, así que sólo interactúan los primeros vecinos, de menor alcance que aquellas presentes en cristales reales. En la figura 1-12 muestro la forma de la energía potencial para burbujas de diferentes diámetros.
4. La masa de las burbujas - Es difícil reproducir los efectos de una energía cinética asociada a una burbuja.

En la figura 1-13 se puede observar una dislocación en un arreglo bidimensional de burbujas. La figura 1-14 muestra dos fronteras de grano formadas por burbujas.

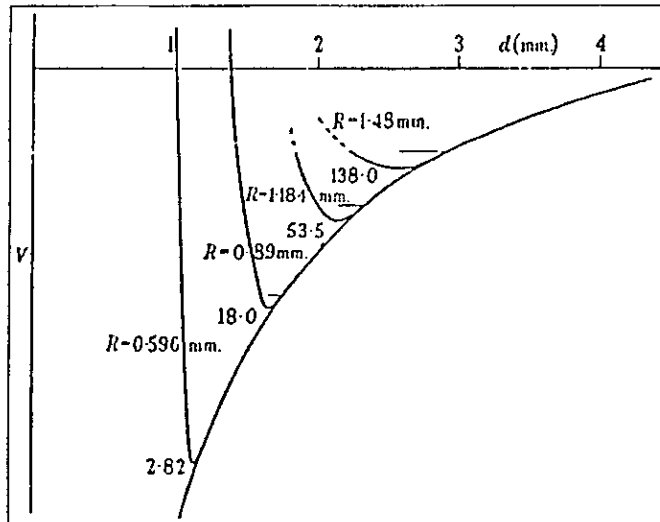


Figura 1-12 Curvas de energía potencial V para burbujas de diferentes radios R y separaciones entre ellas d . [Ish76]

Por último me gustaría agregar la receta original de Bragg y Nye [Bra47] para la solución jabonosa:

ácido oléico	15.2 cm ³	}	mezcla	}	mezcla
agua	50 cm ³				
10% triethanolamina	73 cm ³				
agua	61.8 cm ³				
glicerina pura	164 cm ³				

(1.6)

Para una descripción completa de la forma de generar burbujas y fotografiarlas sugiero el texto de Ishida [Ish76].

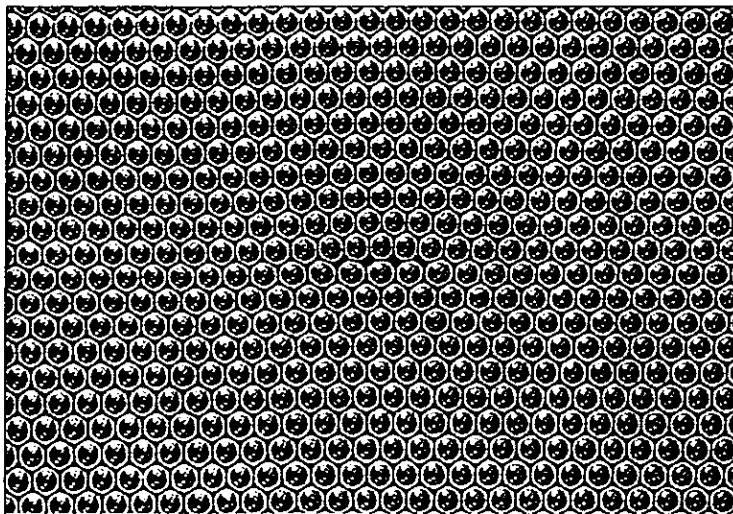


Figura 1-13 Una dislocación en una configuración bidimensional de burbujas. Es posible verla más fácilmente rotando la página unos 30° sobre su plano y observar a bajo ángulo. [Kit86]

1.3.9 Modelo de la "O-Lattice"

Lo siguiente es del texto de Michael Finnis y Manfred Rühle [Fin92, pág 541].

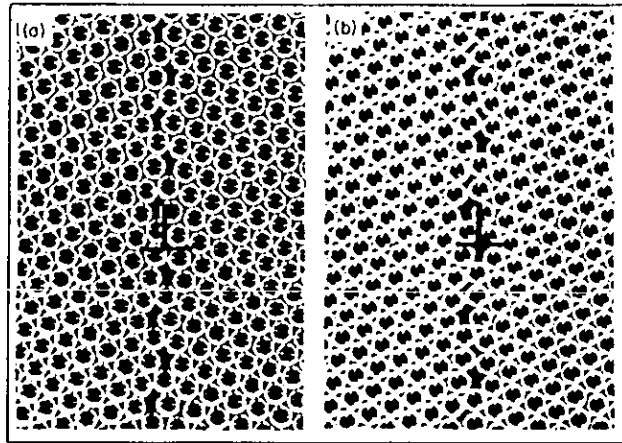


Figura 1-14 Dislocaciones en fronteras a) $\Sigma = 13$ y b) $\Sigma = 31$. Se pueden observar las estructuras de las dislocaciones. [Kit86]

Bollman propuso en 1970 [Bol74] una generalización del modelo de red de coincidencia CSL. El modelo es llamado de la *O-lattice* (de la *red-O*). La definición de la red-O es independiente de la localización de la interface, al igual que la CSL, aunque al final sólo nos importan los puntos de la red en el plano de la frontera.

Comenzando con dos latíces idénticas interpenetradas A y B que coinciden en un punto que consideramos origen. Supongamos un mapeo lineal homogéneo que lleve A a completa coincidencia con B en todo punto. El mapeo puede ser una rotación pura o alguna otra operación que pueda ser representada por una matriz multiplicando un vector.

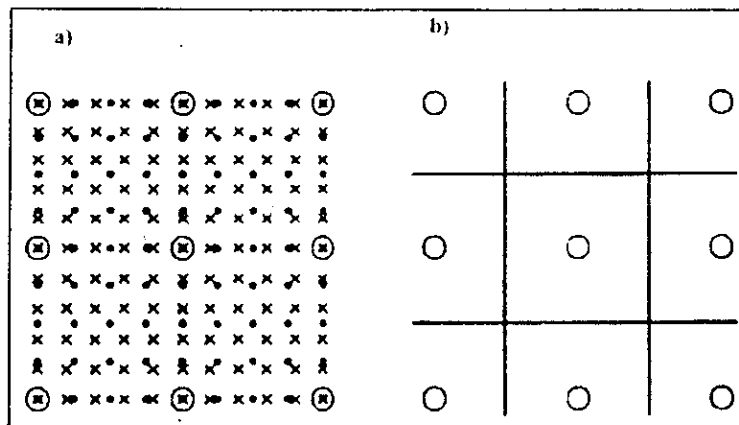


Figura 1-15 Una proyección (001) de la *O-lattice* entre dos redes cúbicas simples con diferente parámetro de red. a) Los átomos de cada cristal están representados por puntos y cruces, los puntos-O son círculos. b) Muestra las líneas medias entre las puntos-O que pueden ser consideradas como dislocaciones. Notese como las dislocaciones corresponden con las regiones de peor encaje [Fin92, pág 541].

Los puntos de coincidencia son completamente equivalentes unos con otros, así que cualquiera de ellos puede ser considerado el origen de la transformación. En general hay aún más puntos que pueden ser tomados como origen del mapeo; el conjunto de todos ellos (puntos O) es la Red-O. Ahora queda claro el nombre O de Origen. Puede haber líneas (líneas-O) y planos de puntos O (planos-O). Por ejemplo cuando la transformación es una rotación hay una línea completa de puntos sobre el eje de la rotación.

En otras palabras, en una red de puntos de coincidencia sólo puntos de las redes pueden ser considerados como de coincidencia. La red-O que propone Bollman extiende las posibilidades permitiendo que cualquier punto del espacio sea un punto de coincidencia. Sólo es necesario que puedan ser considerados como origen para que sean puntos de la red-O. De esta forma es claro que la red de puntos de coincidencia (CSL) es un subconjunto de la red-O (*O-lattice*).

Los puntos O son aquellos en los que las coordenadas internas de un punto en una celda unidad de la red A son los mismos que en una celda unidad de la red B, considerando los orígenes y ejes de cada red respectivamente.

Una ventaja conceptual de la red O sobre la CSL es que el conjunto de puntos O cambia en forma continua

mientras se rota la red A respecto de la B, mientras que en CSL los puntos de coincidencia aparecen y desaparecen. Inclusive hay puntos O sin coincidencia atómica entre las redes. Esto tiene un efecto importante: la red O se puede aplicar a fronteras heterofase y cuando los cristales son incommensurados. Por ejemplo, dos redes cúbicas con la misma orientación pero con ligeramente diferentes parámetros de red. Una puede ser transformada en la otra por una expansión homogénea. Si las caras (100) están en contacto, se forma una red cuadrada de puntos O en los puntos en que la red O interseca la frontera.

Geométricamente se consideran a las líneas que bisectan el arreglo de puntos O en la frontera como las dislocaciones que concentran el desajuste. Por lo tanto hay una relación cercana entre la red-O y la teoría geométrica de dislocaciones en las fronteras [Bol74]. Bollman hace notar que estas dislocaciones son una herramienta matemática y no necesariamente corresponden con dislocaciones físicas observables en estructuras atómicas.

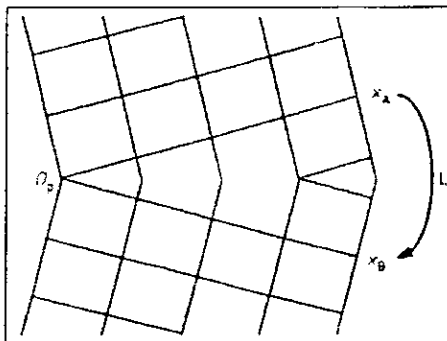


Figura 1-16 Ilustración de la definición de un punto-O. Si un punto x_A de la latiz es transformado a un punto x_B por la transformación lineal L actuando sobre O_P , entonces O_P es un punto-O [Fin92, pág 541].

Una crítica al concepto de la red O parte del hecho de que para una orientación dada entre A y B, la transformación que lleva A a coincidencia con B no es representable por una única matriz. Dado que la red O es por definición una propiedad de la transformación y no de las orientaciones, no debe ser una sorpresa que no sea única. Es esta no-unicidad por la que uno debe tener mucho cuidado al vincular un significado físico en términos de dislocaciones observables a las dilocaciones matemáticas definidas por la red O.

1.3.10 Red DSC

Lo siguiente es del texto de Michael Finnis y Manfred Rühle [Fin92, pág 542].

Otra red que es de importancia al estudiar dislocaciones aisladas y escalones en fronteras es llamada la red DSC, por el inglés de *Displacement Shift Complete*. Esta definida como la red más gruesa que incluya como subredes a A y B (en una orientación en que haya alguna coincidencia). Cualquier vector que una un punto de la red del cristal A con un punto de la red de los cristales A o B, y vice versa, es también un vector de la red DSC.

Por lo tanto, si cualquier red es trasladada por un vector de la DSC, el patrón completo de las redes interpenetradas y la CSL es invariante.

La red DSC es única dadas las orientaciones de A y B y es útil para predecir o explicar las dislocaciones físicas (observables) en fronteras.

1.4 Características

"Ese es el problema con la ciencia. Un monton de gente tratando de describir un mundo de maravilla inimaginable."
- Calvin (& Hobbes)

1.4.1 Orientación de los granos

Es bien conocido que los granos de metales y compuestos rara vez están orientados al azar, y por lo general tienen una textura global. El problema es decidir si esta textura es el resultado de la existencia preferida de fronteras de grano especiales o de algún proceso que controla las orientaciones de los granos directamente [Pum76].

La razón para la existencia preferida de fronteras de grano especiales en un policristal depende del momento en la historia termomecánica del material en que se formaron las fronteras. Si se originaron en la nucleación de los granos, es más probable que sus desorientaciones dependan de las energías relativas de las fronteras que si resultan del encuentro de granos en crecimiento. En este último caso, las movilidades de la frontera pueden ser más determinantes en la distribución final.

1.4.2 Fronteras especiales

Hay mucha evidencia experimental de que las propiedades de las fronteras varían tanto con la desorientación como con el plano de la frontera [Cle72]. En particular, hay algunas desorientaciones especiales en las que las fronteras se comportan como si tuvieran menor energía que el resto, y son llamadas *fronteras especiales* (o *singulares*). Para explicar lo anterior los modelos de frontera deben mostrar alguna relación entre la estructura y parámetros cristalográficos. Todas aquellas fronteras que no son consideradas especiales se les llama *fronteras aleatorias* (o *generales*) (random boundaries en inglés).

Casi toda la evidencia de la existencia de fronteras de grano especiales es indirecta y su existencia ha sido deducida de sus propiedades [Pum76]. sin embargo existen algunas evidencias directas con mediciones de energía y observaciones de la estructura. La mayor parte de los experimentos se realizan en bicristales con desorientaciones predeterminadas, lo que significa que los resultados no necesariamente son los mismos que en policristales.

1.4.3 Propiedades

A continuación describo algunas de las propiedades importantes de las fronteras de grano.

1. **Energía.** Todos los modelos descritos en la sección 1.3 predicen una dependencia de la energía de la frontera de grano en parámetros cristalográficos, con máximos y mínimos para ciertas desorientaciones específicas [Pum76]. Según el modelo CSL [Bra64] el plano de la frontera tiene un efecto importante en la energía de la frontera, porque estructuras de alta coincidencia deben ocurrir en planos de frontera específicos.

Diferentes cálculos computacionales de la energía de fronteras de inclinación a $0^\circ K$ muestran mínimos, que a su vez dependen del ángulo de rotación [Has71]. De la figura 1-17 Pumphrey [Pum76] concluye que una corta periodicidad en el plano de la frontera es un buen criterio para baja energía. El modelo CSL asume que la profundidad de los mínimos decrece cuando Σ crece, lo que no se observa experimentalmente.

Es difícil comparar las mediciones directas de la energía y los cálculos, porque los cálculos acostumbra realizarse a $0^\circ K$, mientras que las mediciones se hacen necesariamente a mayores temperaturas, y porque si la entropía de las fronteras especiales es menor que aquella de las fronteras aleatorias, las diferencias de energía entre fronteras especiales y las aleatorias debe disminuir al aumentar la temperatura.

Si se permite que dos granos roten libremente, lo harán hasta llegar a la desorientación de energía mínima. La figura 1-24 muestra la frecuencia de las desorientaciones observadas para humo de MgO. Los autores [Cha70] muestran claramente que hay CSL's con ejes de desorientación $\{100\}$ son favorecidas. No sólo son favorecidas las fronteras con alta densidad de puntos de coincidencia $\Sigma = 5, 13, 25$ ($\theta = 36.9^\circ, 22.6^\circ$ y 16.3°) sino también aquellas con $\Sigma = 157, 85, 101, 41$ y 37 ($\theta = 8.2^\circ, 8.8^\circ, 11.4^\circ, 12.7^\circ$ y 19.0°).

En un policristal normal, los granos no pueden rotar libremente para llegar a posiciones de mínima energía. Sin embargo, el plano de la frontera es libre de migrar, aunque este movimiento está limitado porque aumenta el área de la frontera, incrementando la energía.

2. **Estructura.** Todos los modelos de la sección 1.3 predicen que a escala atómica las fronteras especiales son ordenadas y tienen estructuras atómicas de alto empaque. A una escala ligeramente mayor que la atómica

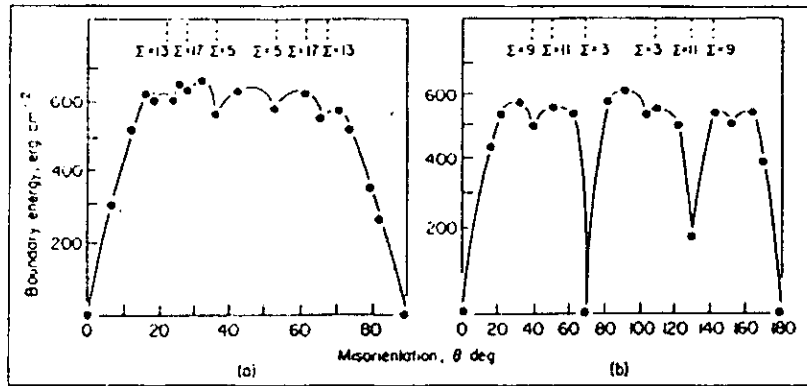


Figura 1-17 Energía de la frontera calculada a 0°K para fronteras de inclinación simétricas de aluminio: a) Eje de rotación [100]; b) Eje de rotación [110]. La energía muestra un mínimo en 70.5° y 129.5°, correspondientes a fronteras CSL con $\Sigma = 3$ y $\Sigma = 11$ [Has71].

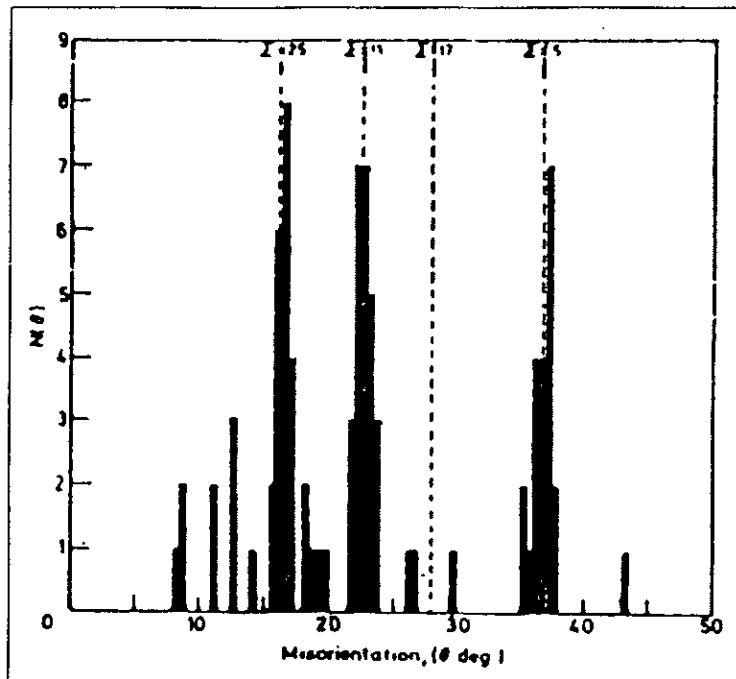


Figura 1-18 Histograma del número de fronteras [100] en humo de MgO como función del ángulo θ [Cha70].

los modelos CSL de unidades estructurales y de casi coincidencia predicen escalones en la interface que siguen planos de relativamente baja energía [Pum76].

Al desviarse una frontera de la desorientación de las fronteras especiales, los modelos de CSL y de planos coincidentes predicen arreglos de defectos para acomodar las regiones de mayor energía (figura 1-7) [Pum76].

La microscopía de campo de iones de fronteras de grano muestra que las regiones de frontera son muy delgadas, con un ancho típico de 2 a 3 átomos [Bra61].

De observaciones de redes de dislocaciones se puede deducir la existencia de fronteras CSL con $\Sigma = 3, 5, 9, 13a, 17a$ (y quizás también de $\Sigma = 11$ y 25) con energías menores que las de fronteras aleatorias. No es clara la relación entre Σ y la magnitud de los máximos y mínimos de la energía.

Es posible que los efectos electrónicos sean más importantes que una coincidencia puramente geométrica. Esta teoría explica con las diferencias electrónicas las grandes variaciones en la energía de fronteras en diferentes metales.

Según [Pum76] hay considerable evidencia de que las fronteras conservan estructuras de relativamente poca energía cuando tienen desorientaciones ligeramente diferentes a las de coincidencia geométrica. Sin embargo, es necesario un criterio para determinar la máxima desviación a la cual una frontera sigue siendo especial. Las mediciones experimentales de propiedades de fronteras ligeramente diferentes de las especiales es difícil, porque no hay un método para preparar muestras de esta forma y no es fácil distinguirlas en microscopía porque el contraste de las dislocaciones disminuye al aumentar su densidad.

La máxima desviación aceptable de una frontera aleatoria de una desorientación especial esta dada por [Pum76]:

$$\rho = \frac{b}{d_{\min}} \quad (1.7)$$

donde b es el vector de Burgers de la dislocación y d_{\min} la mínima distancia entre dislocaciones (que debe ser del orden de la periodicidad). Aún cuando d_{\min} está relacionada con Σ , la relación precisa es difícil de determinar. Entre las propuestas están la de [War74] $\rho = \rho_0 \Sigma^{-2/3}$ y la de [Ish73] $\rho = \rho_0 \Sigma^{-1}$ (donde ρ_0 es constante para todas las CSL's).

La figura 1-19 muestra la separación entre dislocaciones observada en bicristales de oro con frontera de rotación, como función de la desorientación [Sch70].

Pumphrey explica como ni el modelo CSL ni el de planos coincidentes pueden extenderse por medio de dislocaciones para cubrir todas las fronteras [Pum76], y a pesar de la fuerte discusión al respecto de cuantas pueden explicar siempre queda un rango importante que no pueden.

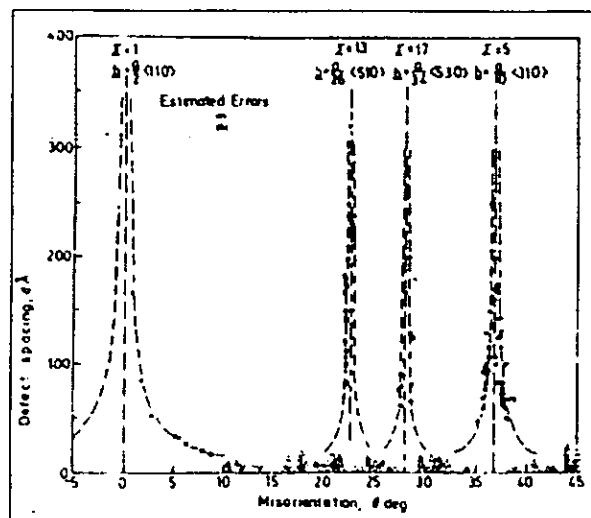


Figura 1-19 Separación entre las dislocaciones observadas en bicristales de oro con fronteras [001] como función de la desorientación. Las curvas punteadas son cálculos según $d = |b|/\Delta\theta$, donde $\Delta\theta$ es la desviación de la desorientación CSL. Los triángulos pequeños (en la base) indican fronteras en que no se detectaron redes de dislocaciones [Sch70]. Notese que las separaciones llegan a ser de cientos de Å.

3. **Segregación.** Todos los materiales metálicos reales tienen alguna proporción de impurezas, por lo que la interacción entre ellos y las fronteras de grano tiene importancia práctica. Se distinguen dos tipos de interacciones:
 - (a) Cuando la fuerza que la propicia es solamente la de interacción entre los átomos y la frontera (segregación de equilibrio).
 - (b) Cuando hay alguna fuerza adicional induciendo el movimiento de átomos disueltos (segregación fuera de equilibrio). Por ejemplo, cuando hay un gradiente de concentración o interacciones entre los átomos disueltos y vacancias.

La evidencia experimental indica que la concentración de impurezas en la frontera es entre 5 y 10 veces mayor que en el cristal. Sin embargo, estimaciones del ancho de la zona de segregación varían desde 4 Å (impurezas en tungsteno) hasta los 900 Å de las impurezas de oxígeno en irido. [Pum76]. Un ancho grande indica que la segregación no es de equilibrio. Se ha reportado una correlación entre alta segregación de impurezas y buen encaje en la frontera [For67].

La energía de fronteras de grano en fierro y cobre disminuye 50% con la adición de 0.8% de Sn [Inm63] y 1% Sb respectivamente [Hon72].

4. **Corrosión.** Las fronteras de grano son más susceptibles de ataque químico dada su mayor energía libre con respecto al cristal perfecto. Por lo tanto fronteras especiales de energía particularmente baja serán también más resistentes a ataque químico.

La segregación hacia las fronteras modifica su composición química, favoreciendo la corrosión.

La velocidad de corrosión en fronteras de alto ángulo cambia mucho más rápido que la energía, lo que parece significar que no es suficiente un criterio energético para predecir efectos de corrosión en fronteras [Pum76].

5. **Difusión.** La difusión en metales es mayor a lo largo de las relativamente desordenadas fronteras de grano que en el cristal perfecto. La solución de las leyes de difusión de Fick requiere suponer que la frontera tiene un ancho δ [Pum76], de lo contrario obtenemos $D\delta$ en lugar de D (D es la difusividad).

Los modelos capaces de hacer predicciones de difusión lo logran solamente para fronteras de bajo ángulo haciendo suposiciones de las dislocaciones que la componen. Sin embargo, al aumentar la densidad de dislocaciones en fronteras de alto ángulo dejan de ser aplicables [Tur54]. Los modelos anteriores al GCSN excluyen la posibilidad de diferentes propiedades de difusión para fronteras especiales.

Experimentalmente se observa una marcada anisotropía en la difusividad, tal y como muestra la figura 1-20. Aún no ha podido explicarse la anisotropía en la difusión con los modelos existentes, que incluso parecen predecir lo opuesto.

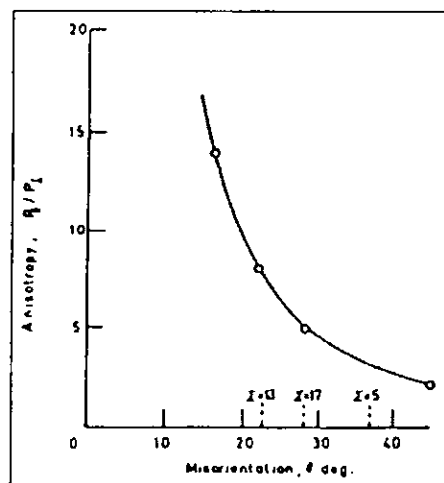


Figura 1-20 Variación de la anisotropía ($P_{||}/P_{\perp}$) de la difusión en fronteras de grano de inclinación [100] de plata a 450°C como función de la desorientación [Hof56].

Los modelos de la sección 1.3 predicen mínimos en la difusividad para orientaciones CSL porque estas tienen mayor empaque que las fronteras aleatorias. La evidencia experimental de lo anterior es ambigua [Pum76].

La difusión tiene un efecto importante en las deformaciones al influir en el proceso de deslizamiento de una frontera de grano. Una frontera sobre la que la difusión sea baja puede ser núcleo de una fractura y dar inicio a la falla del material [Pum76].

6. **Precipitación.** La precipitación depende fuertemente de la difusión, pero también de la estructura de la frontera. La precipitación de estaño en policristales de plomo a temperatura ambiente es unas 10^6 veces mayor que la difusión en un cristal perfecto [Kit86].

Se sabe también que la precipitación es generalmente mayor en las fronteras de grano, debido en parte a la mayor energía libre de estas regiones

También se observan fronteras con valores especiales para con esta propiedad, y no ha sido encontrada una variación sistemática de la densidad de precipitado como función de la desorientación (figura 1-21) [Lec70].

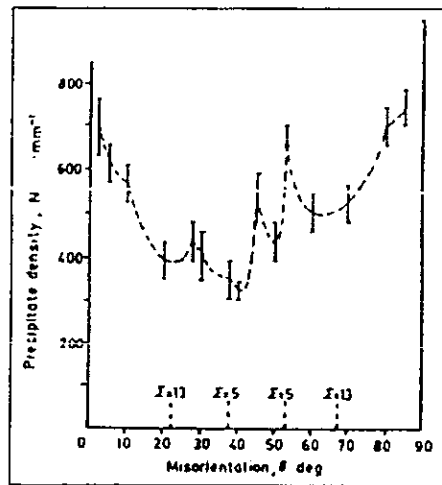


Figura 1-21 Precipitados por mm (N) en bicristales de Aluminio con 3.5% Cobre con frontera de inclinación [100] [Lec70].

7. **Propiedades mecánicas.** Las fronteras de grano tienen efectos considerables en las propiedades mecánicas de los policristales. Se acostumbra dividirlos en efectos a baja y a alta temperatura. A altas temperaturas, cuando la difusión es significativa, hay deslizamiento paralelo al plano de la frontera [Pum76].

De los experimentos es claro que las propiedades de una frontera de grano dependen de sus parámetros cristalográficos, pero en muchos casos, es difícil explicarlo con los modelos de la sección 1.3. Las razones que Pumphrey [Pum76] da para esto son:

- (a) Los modelos son esencialmente para fronteras estacionarias, en metales puros a $0^{\circ}K$. Todos los metales reales tienen impurezas y las propiedades de las fronteras se miden a temperaturas mayores. No es claramente válido relacionar las propiedades dinámicas con propiedades estáticas.
- (b) Algunas propiedades de las fronteras de grano dependen principalmente de la orientación de los granos y no de la estructura de la frontera. Por ejemplo, la propagación de un deslizamiento depende de la orientación del sistema en relación con el eje de tensión: la nucleación de precipitados depende de la interface entre el precipitado y el cristal.

Hay propiedades asociadas con fronteras CSL que aparentemente persisten por varios grados alrededor, mientras que otras fronteras CSL con valores similares de Σ no parecen tener ninguna propiedad especial en absoluto [Pum76].

8. **Dureza.** La fractura quebradiza en sólidos policristalinos ocurre ya sea por división transgranular o por separación intercrystalina [Hon76]. Numerosas observaciones han mostrado que las fronteras de grano son susceptibles de serias debilidades, principalmente por la presencia de elementos o partículas precipitadas que alteren la química de la frontera y la cohesión. Sin embargo, en sistemas cúbicos centrados en la cara pueden ocurrir fracturas transgranulares, lo que significa que las fronteras de grano no necesariamente son caminos de menor dureza [Hon76]. La figura 1-22 es una fotomicrografía de microscopio de electrones por escaneo de una

muestra de cobre (con 0.006% Bi). En la figura se muestra una superficie después de la fractura y se puede observar que sigue un camino intergranular.

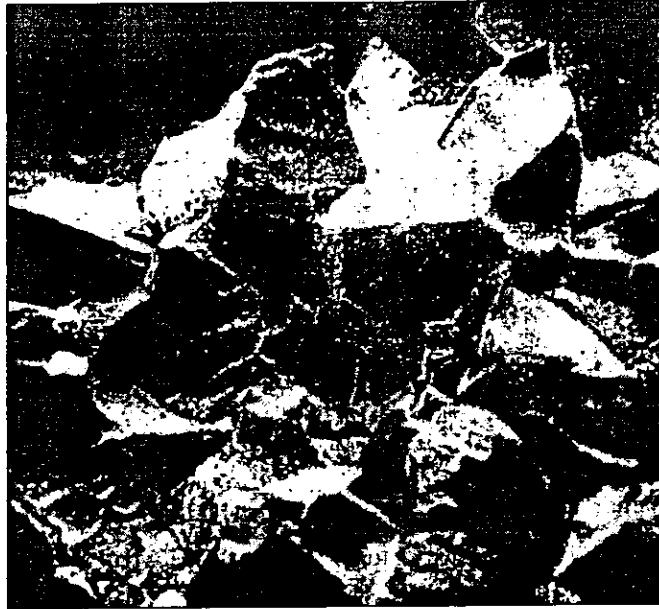


Figura 1-22 Fotomicrografía de microscopio de electrones de la superficie de fractura de un sólido de cobre con 0.006% Bi [Hon76].

9. **Migración.** El fenómeno de migración de una frontera de grano ocurre cuando se somete a una fuerza suficiente para causar movimiento. La migración resultante es tal que reduce la energía libre del sistema [Sim76]. La velocidad con que migra una frontera depende de la temperatura. La presencia de algunos átomos ajenos al cristal, inclusive en concentraciones de unas cuantas partes por millón reduce drásticamente la movilidad. La figura 1-23 muestra la movilidad de una frontera de grano.

La energía de activación de la migración de fronteras de grano es una función de la desorientación, lo que puede modificar las distribuciones de desorientaciones observadas. La figura 1-24 muestra la energía de activación como función de la desorientación para fronteras de inclinación $\{100\}$ [Rut65].

1.5 Crítica de los modelos

El microscopio de campo de iones ha mostrado que la topografía de materiales trabajados en frío es, por lo general, complicada y no sigue las predicciones de ninguna de las teorías [Lob76].

De una serie de mediciones de la energía de fronteras [Pum76] concluye que ni el valor de Σ de la red CSL, ni planos coincidentes ni la periodicidad de la frontera son suficientes para explicar la magnitud de la energía de las fronteras especiales.

Desde el punto de vista de los modelos de la sección 1.3 las razones de migración mayores de las fronteras CSL de alto empaque mostradas en la figura 1-24 son inesperadas [Pum76].

“Es difícil creer que la estructura de las fronteras especiales predicha por los modelos de la sección 1.3 es el único factor que controla la corrosión en la frontera de grano. Esta claro que en un policristal la velocidad de ataque químico varía tanto con la desorientación como con la dirección del plano de la frontera. Por el momento (1976) no es posible predecir de los modelos cuales fronteras serán más resistentes” [Pum76].

Los modelos de la sección 1.3 tienen grandes logros en la descripción de la estructura de fronteras de grano especiales, sin embargo tienen dificultades importantes al hacer predicciones de las propiedades de la frontera, de su efecto en las propiedades del material y de la variación de estas con el ángulo de desorientación.

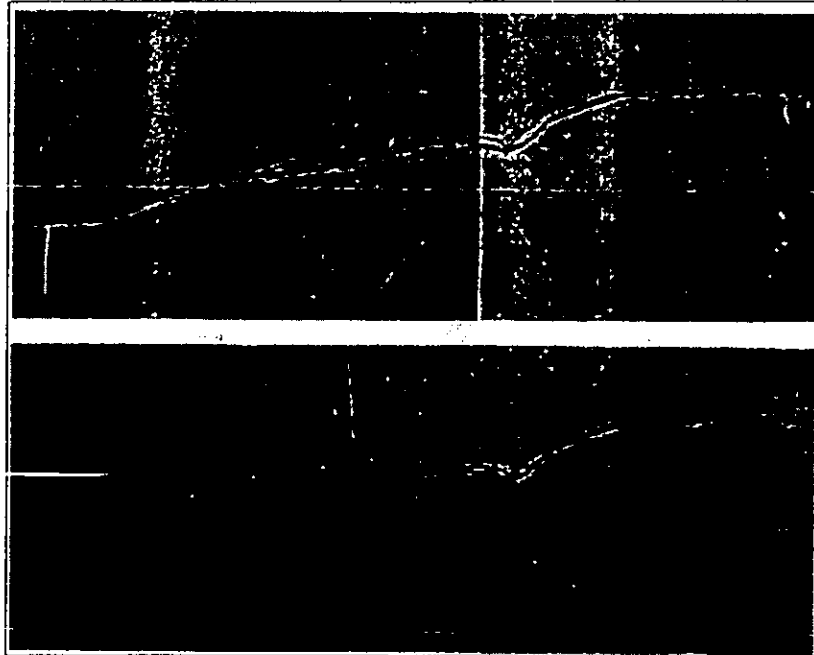


Figura 1-23 Movilidad de una frontera de grano. La frontera es la línea vertical, mientras que la línea horizontal irregular es un pequeño escalon en la superficie que sirve como referencia. El cristal está anclado a la izquierda y a la derecha se sujeta a una fuerza normal al plano de la página. Arriba la posición original de la frontera. Abajo, desplazada 0.4 mm. [Kit86]

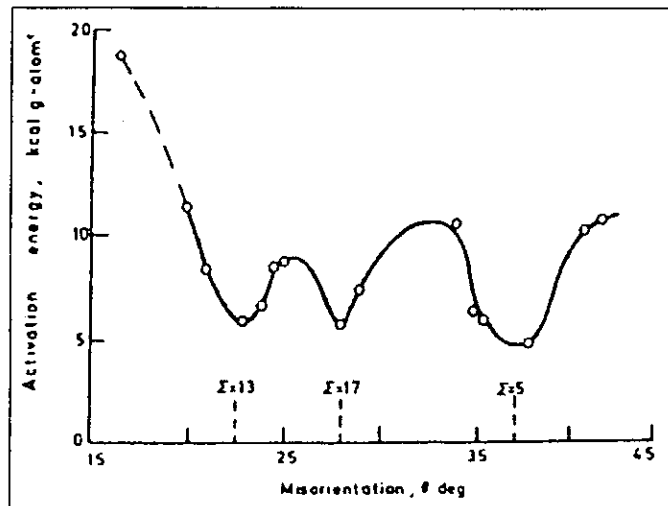


Figura 1-24 Energías de activación para la movilidad de una frontera de grano como función de la desorientación en fronteras inclinación [100] [Rut65].

2. EL MODELO GCSN

“Los científicos tienen terribles modales, excepto cuando se elogia su teoría; entonces uno puede hasta pedirles dinero prestado” Mark Twain [Bri99].

Lo siguiente es del artículo de Romeu, Beltrán, Aragón y Gómez [Rom99].

2.1 La necesidad de un mejor modelo

Los modelos descritos en la sección 1.3 utilizan criterios geométricos (que no toman en cuenta en forma explícita la interacción entre los átomos de la interface [Sut87]) para determinar la estructura atómica de una frontera de grano como función de los grados de libertad que definen la orientación relativa entre los cristales y el plano de la frontera.

A pesar de que casi todos los modelos han sido exitosos al describir las propiedades de fronteras de grano especiales, han fallado en dar su estructura atómica detallada, que ha tenido que ser inferida de simulaciones por computadora. A pesar de que estos cálculos han dado una buena idea de la relación entre la estructura de la frontera y sus propiedades [Wol92], están restringidas por su naturaleza discreta. Por lo tanto el problema general de relacionar la estructura de fronteras de grano arbitrarias (aleatorias), por mucho las más abundantes, ha recibido poca atención. Esto ha sido un factor que detiene el desarrollo de una cristalografía de fronteras de grano.

Según Gifkins [Gif69] un modelo para frontera de grano satisfactorio debe hacer lo siguiente:

1. Debe describir el cambio de una orientación a otra en términos que permitan colocar los átomos en posiciones acordes con las mismas leyes de fuerza que los unen en un cristal perfecto.
2. El modelo debe predecir la correcta dependendencia de la energía libre en la magnitud y orientación de la frontera.
3. Debe dar un ancho para la frontera de grano que no sea mayor a tres diámetros atómicos.
4. El modelo debe ser una base para describir el rango completo de propiedades de manera cuantitativa: migración, deslizamiento, fusión, segregación, máximos de fricción interna, precipitación, etc. (sección 1.4).

Aaron y Booling [Aar72] agregan el volumen libre como una forma de evaluar modelos. (definen el *volumen libre* como $V_{free} = (\partial V / \partial A)_{T,P,n,x}$) aunque agregan: “En principio V_{free} es medible experimentalmente, pero es dudoso que puedan realizarse experimentos inequívocos”.

2.2 Introducción

Para ayudar a salvar las limitaciones atrás expuestas se ha introducido un modelo estructural llamado *Red de sitios de coincidencia generalizados* (GCSN por las siglas en inglés de *Generalized Coincidence Site Network*) [Rom97, [Gom98]. El modelo GCSN generaliza el modelo de sitios de coincidencia (sección 1.3.5) remplazando el criterio de estricta coincidencia por uno de cuasicoincidencia. Esta definición permite el estudio de fronteras de grano usando las herramientas matemáticas desarrolladas en años recientes para cuasicristales, unificando ambos campos. El criterio de cuasicoincidencia no es en sí mismo lo que hace al GCSN especial y útil. La clave del modelo esta en la manera en que se construyen las fronteras de grano (entre bicristales). Contraria a la manera usual de simplemente yuxtaponer dos rebanadas de cristal, los bicristales GCSN tienen un volumen de frontera de grano diferente, con átomos en posiciones determinadas por puntos de cuasicoincidencia, que se asume minimizan la energía elástica en la interface, una hipótesis a verificar *a posteriori* tanto empírica como numéricamente.

A pesar de que el modelo GCSN contiene suposiciones implícitas de las relajaciones en la interface, y no puede considerarse como puramente geométrico [Sut87], no puede predecir cuales fronteras de grano tendrán comparativamente menor energía. El modelo suministra una configuración atómica inicial para cada frontera que se espera sea cercana a la de un mínimo local para la mayoría de los potenciales (centrales) usados para modelar metales FCC (sección 7). Para determinar cuales fronteras son favorecidas energéticamente es necesario realizar cálculos atómicos en bicristales GCSN usando algún potencial adecuado (ver sección 3).

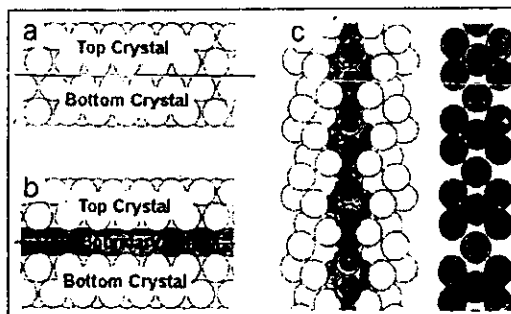


Figura 2-1 a) Construcción común de un bicristal con dos rebanadas de cristal yuxtapuestas. b) Bicristal GCSN con una región de frontera (en gris). c) Izquierda, bicristal de inclinación con los átomos de la frontera destacados en gris; derecha, vista lateral de los átomos de la frontera [Rom99].

El modelo ha logrado reproducir en forma cuantitativa importantes observaciones experimentales acerca de la estructura de las redes de dislocaciones primarias y secundarias en fronteras de grano. También ha logrado introducir una definición de especialidad basada en un mínimo contenido de dislocaciones y reportado otras predicciones fundamentales concernientes a la estructura atómica detallada de fronteras especiales (también llamadas delimitantes [Vit80]) y alcatorias (o intermedias [Vit80]), así como del rango angular sobre el que la estructura delimitante se extiende.

2.3 Detalles

La siguiente es una breve descripción no matemática del modelo GCSN, para un tratamiento más riguroso se pueden consultar los artículos [Gom98] [Ara97].

El modelo de sitios de coincidencia [Kro49] considera puntos comunes de ambas latíces como puntos de buen ajuste, y asume que las fronteras especiales son aquellas en las que el recíproco de la densidad de sitios de coincidencia (Σ) es bajo. El modelo GCSN también está definido como un conjunto de puntos de buen ajuste entre dos redes, pero aquí el requerimiento de coincidencia estricta se reemplaza por uno de cuasicoincidencia definido así: Dos puntos (uno en cada red) son cuasicoincidentes si ambos se encuentran en la intersección de sus respectivas celdas de Wigner-Seitz. Una definición más intuitiva es que dos puntos, o sitios son cuasicoincidentes si cada uno es el único vecino más cercano del otro, y su distancia es menor que la suma de los radios atómicos de los átomos que ocupan los sitios. El GCSN se define entonces como el conjunto de puntos medios, o promedio, entre los pares cuasicoincidentes. Hay que hacer notar que, a pesar de haber usado el término latíz, la definición de GCSN no requiere que las estructuras interpenetradas sean periódicas, permitiendo la construcción de interfaces cristal-cuasicristal y cuasicristal-cuasicristal.

Dado que el criterio de cuasicoincidencia es obviamente satisfecho por puntos estrictamente coincidentes, el modelo GCSN contiene la CSL asociada, pero más importante es que contiene también su decoración (figura 2-4). Por lo tanto para latíces decoradas el GCSN produce la estructura atómica detallada de la frontera incluyendo las redes de dislocaciones primarias y secundarias. Notese que, a diferencia de la CSL, GCSN sigue bien definido cuando la tangente de la desorientación es irracional. Sin embargo, en este caso el GCSN es una red aperiódica en lugar de una latíz [Rom97] [Ara97].

Un bicristal GCSN se construye de la siguiente manera:

1. Dos latíces arbitrarias se llevan a tener un origen común, y una de ellas es rotada sobre un par eje/ángulo. Opcionalmente una de las latíces puede ser desplazada por un vector arbitrario.

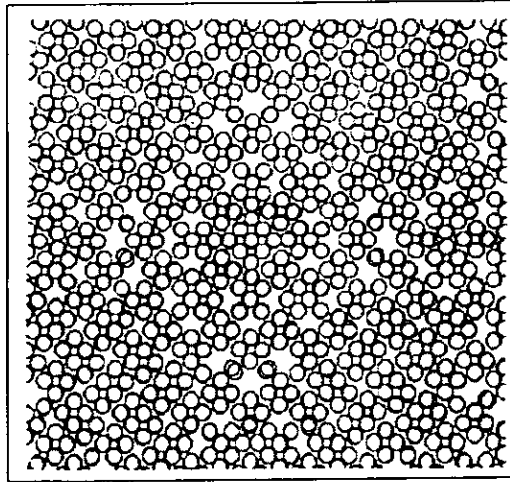


Figura 2-2 Graficas superpuestas de átomos en una frontera aleatoria ($\theta = 23.62^\circ$) antes y después de una relajación estática llevada a cabo usando el potencial de Níquel de Bristowe y Crocker [Rom99].

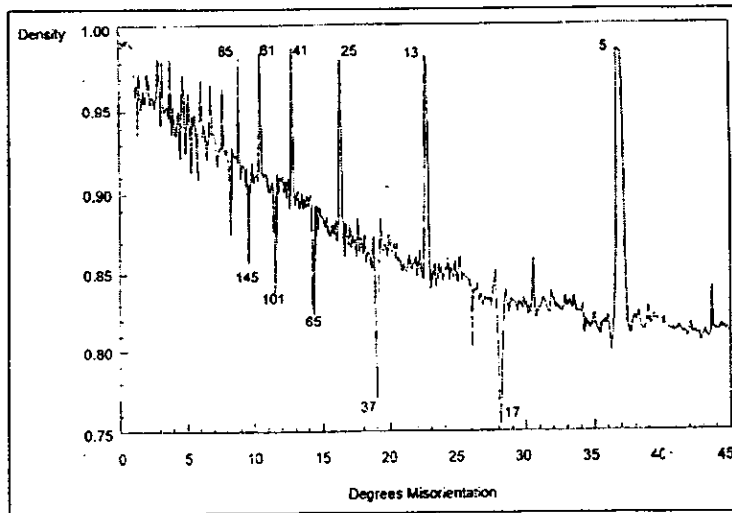


Figura 2-3 Gráfica normalizada de la densidad atómica vs ángulo de desorientación en grados para fronteras de rotación fcc [100] calculadas cada centésimo de grado. Los números cerca de los picos identifican el valor de la Σ correspondiente [Rom99].

2. Una rebanada de la frontera de espesor diferente de cero se define pasando por el origen por conveniencia (el grosor de la frontera de grano se fija normalmente en una distancia interplanar).
3. Todos los puntos de las latíces dentro de la frontera se remplazan por el GCSN según describí antes.
4. Los puntos restantes de la latíz 1 por arriba de la frontera se eliminan. Igualmente se hace con los de la latíz 2 que restan por debajo de la frontera. Se dejan entonces solamente los puntos que no han sido usados para generar un punto GCSN en la frontera.

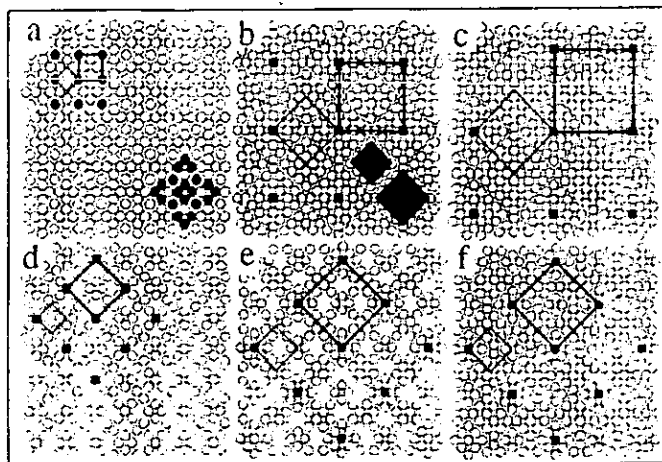


Figura 2-4 a) - e) Estructuras atómicas de $\Sigma 5$, $\Sigma 41$, $\Sigma 61$, $\Sigma 17$ y $\Sigma 37$ respectivamente. f) $\Sigma 37$ con los sitios de vacancias removidos. Círculos abiertos: átomos normales de la frontera de grano. Círculos gris oscuro: nodos CSL. Líneas delgadas: Celda unidad de la red O. Líneas gruesas: celda unidad CSL. Círculos claros en $\Sigma 5$ marcan patrones atómicos también presentes en $\Sigma 17$. Cuadros oscuros en $\Sigma 41$ ($n = 4$) destacan dominios atómicos con cuatro y cinco átomos en la orilla [Rom99].

En el caso específico de fronteras de rotación con ejes de rotación de bajo índice, todos los átomos de la frontera se encuentran en un sólo plano (figuras 2-1(b) y 2-3). Por otro lado, cuando la normal a la frontera no es un plano cristalográfico de bajo índice, la zona de la frontera pre relajada puede tener grosor finito, como muestra la figura 2-1(c).

Se usa la frontera de rotación FCC [100] para hacer cálculos de las predicciones del modelo GCSN porque es el sistema bidimensional más simple que contiene todas las características pertinentes [Rom99]. Sin embargo, el modelo también puede ser aplicado tal y como está al estudio de interfaces cristalinas y no cristalinas, y en general a cualquier sistema para el que se pueda construir una frontera bidimensional cómo se describe atrás. El modelo GCSN puede ser usado para estudiar cristales iónicos, donde la introducción de vacancias en las fronteras ha mostrado reducir la energía, para determinar la localización de tales vacancias. El modelo GCSN propone que cuando iones cuasicoincidentes tengan la misma carga sean remplazados por un ión de esa carga, y cuando tienen cargas opuestas por una vacancia.

2.4 Diferencias con otros modelos

Una diferencia esencial entre este modelo y el resto (sección 1.3) es que en los anteriores dos rebanadas de cristal se yuxtaponen a cada lado de la frontera como muestra la figura 2-1(a). La frontera pre-relajada en el bicristal GCSN contiene sitios atómicos que, excepto por las coincidencias, no pertenecen a ningún cristal [Bol82].

Esta diferencia es crucial y se encuentra en el corazón del modelo. En un bicristal GCSN, cada átomo de la frontera se encuentra en el punto intermedio entre dos sitios cuasicoincidentes (átomos sobrepuestos) de las redes originales. El punto medio es una posición de compromiso entre las ligeramente diferentes posiciones que ocuparía cada átomo del cristal en su propia latíz de no haber interface. Esto causa que átomos en la frontera actúen como un colchón elástico que permite una transición suave entre los dos cristales adyacentes. Como resultado, cuando se

relaja un bicristal GCSN los átomos en la frontera y los de los cristales cambian sus posiciones sólo ligeramente (2-2), dejando la estructura de la frontera de grano practicamente inalterada.

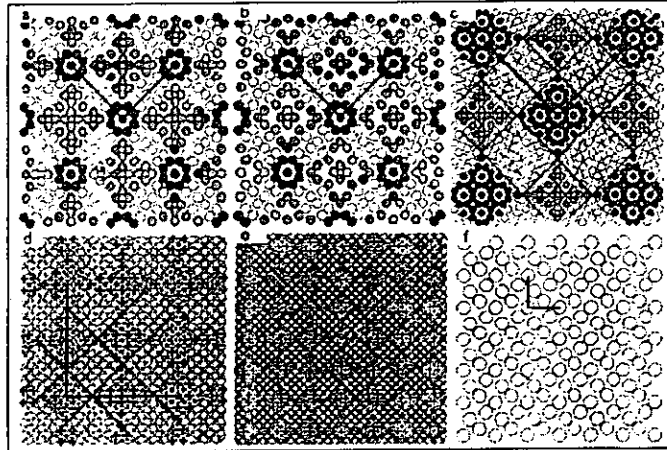


Figura 2-5 Algunas fronteras de grano asociadas a $\Sigma 5$: a) $\Sigma 53$ sin vacancias, b) $\Sigma 53$ con vacancias, c) $\Sigma 185$, d) y e) diagramas de líneas de $\Sigma 461$ y $\Sigma 1189$ con líneas que unen átomos separados una unidad de red o menos. f) $\Sigma 5$ en el estado traslacional tipo CSL. Átomos grises: dominios atómicos en estado traslacional tipo 2. Líneas gruesas: vectores CSL. Líneas delgadas: Posición promedio de las dislocaciones parciales secundarias de frontera de grano.

Según lo anterior, el modelo GCSN debe proveer una buena representación de la estructura de una frontera de grano en cualquier sistema para el que los mecanismos de minimización de energía GCSN sean válidos (presumiblemente la mayoría de los metales con configuraciones de alto empaque (sección 5.1.1)), o cualquier sistema cuya estructura pueda ser modelada por un potencial central (sección 7). Sin embargo, hay que reiterar que la evaluación final de la validez de esta afirmación y la determinación de que sistemas quedan bien descritos por el modelo requiere la comparación de sus predicciones con observaciones experimentales por medio de simulaciones como las de esta tesis.

2.5 Fronteras especiales (singulares)

Para todos los modelos que el modelo GCSN describe adecuadamente, permite estudiar las propiedades estructurales básicas de sus fronteras de grano, sin necesidad de recurrir a cálculos computacionales, haciendo posible estudios cristalográficos de fronteras arbitrariamente grandes sobre rangos angulares completos.

Un ejemplo de cálculo útil realizable que ahora puede ser hecho sobre todo el campo angular, la figura 2-3 muestra una gráfica de la densidad atómica en una frontera de grano contra el ángulo de desorientación para fronteras de rotación FCC [100], calculada con una precisión de un centésimo de grado. La figura describe el número de sitios (pares cuasicoincidentes) en una sección circular del plano de una frontera de grano (el radio del círculo es de 30 separaciones de red).

La gráfica puede ser normalizada para representar cualquier estructura cúbica. En la gráfica se observan picos y hendiduras delgadas en desorientaciones específicas. Ellos son una medida de un buen encaje entre los planos rotados. Resulta que tanto los picos como las hendiduras reproducen las orientaciones de fronteras especiales de baja energía de sistemas reales. Evidencia de esto es que coinciden con las orientaciones identificadas como más probables experimentalmente (ver figura 1-18).

Según Ranganathan [Ran66], una frontera CSL se genera cuando una latíz se rota respecto a la otra sobre un eje $\langle h, k, l \rangle$ por el ángulo

$$\theta = 2 \tan^{-1} \left(\sqrt{N} \frac{y}{x} \right), \tag{2.1}$$

con y y x enteros primos relativos, $N = h^2 + k^2 + l^2$, y $\Sigma = x^2 + Ny^2$ (dividida por dos hasta que Σ sea impar). Usando esta definición, los picos principales de la figura 2-3 ($N = 1$) se encuentran en ángulos θ_n para los cuales $y = 1$ y $x = 2n + n$, mientras que los mínimos ocurren cuando $x = 2n$, $n = 1, 2, \dots$. Por lo tanto todas las orientaciones

especiales corresponden con fronteras de coincidencia para las que $y = 1$ y $\Sigma_n = 2n^2 + 2n + 1$ para los picos y $\Sigma_n = 4n^2 + 1$ para las hendiduras.

En el artículo [Rom99] se puede encontrar una discusión de la estructura de las fronteras, de sus dominios y de las redes de dislocaciones primarias. Romeu *et al* exponen como el modelo GCSN reproduce los resultados experimentales existentes y aglutina los resultados exitosos del resto de los modelos.

2.5.1 Definición de especialidad

Uno de los resultados más importantes del modelo GCSN es la posibilidad de dar una definición precisa de especialidad (que en el modelo es llamada *singularidad*): Aquellas fronteras con $y = 1$ que por tener el menor valor posible de y tienen el mínimo contenido de dislocaciones por celda CSL, una. Cuando $y > 1$ aparece una red de dislocaciones secundarias en la frontera, probablemente aumentando su energía.

Este criterio de especialidad es consistente con observaciones experimentales y se basa enteramente en consideraciones físicas (mínimo contenido de dislocaciones) y numéricas ($y = 1$).

Las vacancias pueden ser llenadas repitiendo el proceso de emparejamiento por cuasicoincidencia usando sólo los átomos que no fueron apareados en la primera iteración o dejando los átomos huérfanos como parte de la frontera. La figura 2-4(f) muestra el resultado de tal operación, y puede ser comparada con la frontera con vacancias de la figura 2-4(e) (ver sección 3).

2.6 Fronteras aleatorias (generales)

Hemos utilizado el modelo GCSN para definir en forma más precisa las fronteras especiales (singulares), para luego encontrar que el conjunto de fronteras especiales depende de la propiedad a la que uno se refiera.

Las fronteras que no son especiales son llamadas aleatorias (o generales). El modelo GCSN asocia a cualquier frontera de este tipo una frontera singular, prediciendo además que la estructura de la frontera de grano arbitraria esta compuesta de una mezcla de unidades estructurales de la frontera especial asociada en diferentes estados traslacionales.

2.6.1 Frontera especial asociada

Para encontrar la frontera asociada a una frontera arbitraria θ_x , Romeu *et al* encuentran una secuencia de fronteras intermedias $\{\Sigma_{x,c,p}\}$ en el intervalo (θ_x, θ_{x+c}) entre la frontera especial (delimitante) Σ_{x+c} y Σ_x . Hay una secuencia diferente para cada valor de c , mientras que el parámetro p recorre la secuencia.

Por ejemplo, la figura 2-5 muestra la estructura atómica de una frontera general (no especial) $\Sigma_{53} = \Sigma_{3,1,2}$ localizada en $\theta_{3,1,2} = 31.89^\circ$ entre las fronteras singulares $\Sigma_{17} = \Sigma_4$ y $\Sigma_5 = \Sigma_3$. Esta frontera es el segundo ($p = 2$) elemento de la secuencia $\{\Sigma_{3,1,p}\}$ y esta compuesto de dos conjuntos de dominios atómicos.

Siendo el conjunto de los números racionales denso, cualquier frontera de grano aleatoria con ángulo de desorientación θ_x puede ser aproximada, según la ecuación 2.1, tan bien como queramos por $2 \tan^{-1}(p/q)$ para algunos enteros primos relativos p y q . Notese que $N = 1$ para el caso de fronteras de rotación (100).

Esto significa que, dada una tolerancia, cualquier frontera de grano puede ser pensada como una de coincidencia con

$$\Sigma_{arb} = p^2 + q^2 \quad (2.2)$$

si p y q tienen paridades diferentes, y

$$\Sigma_{arb} = \frac{1}{2}(p^2 + q^2) \quad (2.3)$$

de lo contrario. La clave es expresar la razón $r = p/q$ en términos de enteros x y c como

$$r \equiv r_{x,c,p} \equiv \frac{1}{x + c/p} \quad (2.4)$$

con x definida cómo

$$x = \mathcal{R} \left[\frac{q}{p} \right] = \mathcal{R} \left[1 / \tan \left(\frac{\theta_{arb}}{2} \right) \right], \quad (2.5)$$

$\mathcal{R}\{\cdot\}$ es la función de entero más próximo y c

$$c = q - px. \quad (2.6)$$

El entero positivo x nombra la frontera singular ($y = 1$) con Σ_x en el ángulo de desorientación especial

$$\theta_x = 2 \tan^{-1} \left(\frac{1}{x} \right) \quad (2.7)$$

que se asocia a la frontera arbitraria $\Sigma_{arb} \equiv \Sigma_{x,c,p}$. Si ahora pensamos en escribir el denominador de la ecuación 2.4 cómo $x' = x + c/p$, entonces el término c/p ya no necesita ser racional, y representa un parámetro de desviación de la frontera especial asociada.

3. LO QUE APORTA ESTA TESIS

“El hombre ocasionalmente se tropieza con la verdad, pero usualmente logra levantarse y continuar”,
Winston Churchill [Dic99].

Los modelos de la sección 1.3 hacen descripciones satisfactorias de la estructura de algunas fronteras de grano especiales, prediciendo propiedades de energía, difusividad y otras en correspondencia con los experimentos. Sin embargo fallan al explicar muchas observaciones importantes, por ejemplo las tazas de migración, resistencia mecánica, difusión y otros. Las observaciones muestran la existencia de fronteras especiales, sin embargo con los modelos anteriores no hay una definición clara de lo que es una frontera especial, mucho menos de sus propiedades y el efecto de estas en las propiedades del material completo.

El modelo de red generalizada de sitios de coincidencia GCSN ha mostrado poder predecir satisfactoriamente estructuras y propiedades de fronteras de grano de rotación [100] en metales FCC, pero además hace predicciones que ningún otro modelo había podido hacer, como la estructura en detalle de las redes de dislocaciones (primarias y secundarias). El modelo da una definición precisa de especialidad basada en propiedades físicas y matemáticas, y permite asociar a cualquier frontera aleatoria una frontera especial.

Sin embargo, aún se desconocen muchas de las implicaciones de proponer fronteras GCSN, aunque con lo que se conoce se pueden plantear muchas preguntas interesantes e importantes. por ejemplo:

- ¿Como son las curvas de energía contra desorientación para diferentes materiales?
- ¿Tienen las fronteras con vacancias y baja densidad menor energía que las aleatorias?, ¿si se llenan estas vacancias con átomos de otras especies se reduce aún más la energía?, ¿que implicaciones tiene esta posibilidad en la segregación y el coeficiente de difusión?
- ¿Como varían los coeficientes de difusión con el ángulo de desorientación?, ¿se predicen adecuadamente en los picos y hendiduras?, ¿y la asimetría también?
- ¿Vale la pena usar diferentes definiciones de especialidad respecto a diferentes propiedades?
- ¿Como varían las propiedades en cristales iónicos?
- ¿Como son los patrones de difracción de rayos x de las estructuras generadas por GCSN?

Para responder a estas preguntas y conocer nuevas que puedan surgir es necesario seguir investigando las implicaciones del modelo para diferentes materiales y situaciones.

Hay dificultades experimentales para estudiar las fronteras de grano. Por ejemplo, las de rotación se complican porque el material obstruye la visión directa de la frontera, a diferencia de las de inclinación, en donde los parámetros interesantes de la frontera pueden ser observados lateralmente. La enorme cantidad de y complejidad de los cálculos requeridos para hacer mediciones y predicciones a partir de una GCSN hace del uso de computadoras una necesidad ineludible en el estudio de fronteras de grano.

La naturaleza de las preguntas que queremos resolver sugiere la simulación computacional como la herramienta adecuada, ya que combina ingredientes teóricos y experimentales. Desde sus inicios el estudio de fronteras de grano ha implicado simulaciones. Hoy en día todos los grupos de investigación en el área cuentan con herramientas computacionales para simular. El campo entero de las fronteras de grano descansa, en buena medida, en simulaciones computacionales.

Hay diferentes métodos para hacer simulaciones, cada uno con sus ventajas y desventajas. El departamento de materia condensada del IFUNAM cuenta ya con un programa de simulación por Monte Carlo que ha dado excelentes resultados haciendo las predicciones que se conocen. Sin embargo, tiene limitaciones inherentes y no es capaz de ayudarnos a responder algunas de las preguntas que nos interesan. Para responderlas es más adecuado el método de dinámica molecular.

La dinámica molecular es una herramienta poderosa, y un programa adecuado permite estudiar no sólo fronteras de grano, sino cuasicristales, nanocúmulos, proteínas y muchos otros temas de interés para la física, la química y otras ciencias. Es el método más utilizado para estudiar fronteras de grano. Contar con un programa de este tipo permitiría al departamento de materia condensada estudiar las implicaciones del modelo GCSN y compararlas directamente con las de otros modelos reportadas en la literatura.

El departamento de materia condensada, y en general el IFUNAM requiere entonces de una herramienta de simulación por dinámica molecular y la experiencia necesaria para usarla, mantenerla, extenderla y modificarla.

Mi objetivo es hacer un programa lo más general posible, de tal manera que pueda adaptarse a este tipo de problemas, pero además integrarse con otras de las herramientas computacionales del departamento, cómo por ejemplo el simulador por monte carlo para hacer simulaciones híbridas. No nos interesa hacer un programa pequeño para responder a alguna pregunta específica, porque a largo plazo es mucho más valioso contar con una herramienta general, para responder a una mayor variedad de preguntas, conocidas y por conocer.

No es suficiente el trabajo de una tesis de licenciatura para estudiar a profundidad las implicaciones del modelo GCSN, por lo que van a ser necesarios más estudiantes interesados en el tema. Para simplificar el proceso de aprendizaje y aumentar el aprovechamiento del estudiante el programa debe ser claro y la tesis comprender todo lo necesario para ayudar a que buenos estudiantes de los últimos semestres de la carrera de física con conocimientos básicos de programación, no familiarizados con los temas de frontera de grano y de simulación puedan aprovecharlo. Esto implica la recopilación y redacción de los conceptos básicos tras estas ramas, así como una cuidadosa planeación de la estructura del programa. Ahorrarle el trabajo de describir los conceptos y técnicas importantes y detallar los puntos clave del programa sería complicarle la vida a futuros estudiantes y retrasar el entendimiento del modelo y el aprovechamiento de la herramienta.

En resumen, los objetivos de mi trabajo de tesis son:

1. Construir los elementos centrales de una herramienta computacional general de simulación por dinámica molecular acoplable con otros métodos de simulación.
2. Utilizar dicha herramienta para calcular algunas propiedades de fronteras de grano construídas según el modelo GCSN, a fin de estudiar sus consecuencias y límites de validez.
3. Estudiar tópicos relacionados con la simulación por dinámica molecular que tengan efectos determinantes para el caso de fronteras de grano.
4. Compilar los conceptos principales de las ramas de frontera de grano y simulación en un texto destinado a posibles futuros estudiantes del tema.
5. Describir los trucos y detalles no explícitos en la bibliografía que permiten transmitir al novato las experiencias adquiridas en este trabajo, con la intención de ahorrarle tiempo y esfuerzo.

Parte II

Una solución: dinámica molecular

4. INTRODUCCIÓN A LA DINÁMICA MOLECULAR

“Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it - an intelligence sufficiently vast to submit these data to analysis - it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom: for it, nothing would be uncertain and the future, as the past, would be present to its eyes” [Lap51].

4.1 Simulación vs solución numérica

La solución numérica resuelve problemas de matemáticas, la simulación busca reproducir sistemas (reales o imaginarios). Las soluciones numéricas pueden ser tan exactas como uno quiera, dependiendo de los métodos usados para resolver las ecuaciones. Las simulaciones heredan la aproximación que las ecuaciones fundamentales hacen de los fenómenos naturales, y por lo tanto no pueden ser exactas.

Al encontrar soluciones numéricas lo importante es la solución en si misma, y no que tenga alguna relación con un sistema “real”. Por el contrario, que un sistema real pueda ser aproximado muy bien por soluciones numéricas no deja de ser una casualidad a la que podemos acostumbrarnos o no. De haber una separación entre los mundos filosóficos de la realidad y de las matemáticas, las soluciones numéricas viven del lado de las matemáticas y las simulaciones del lado de la realidad. Por lo mismo sólo pueden existir soluciones numéricas de problemas matemáticos, pero las simulaciones pueden ser de sistemas reales. Siendo las matemáticas la manera mas exitosa que hemos inventado para describir sistemas reales, muchas veces podemos (y casi siempre deseamos) traducir las simulaciones a problemas matemáticos a resolver numericamente.

Aún suponiendo que las ecuaciones fundamentales de la física correspondan con la naturaleza y el mundo “real”, es imposible resolver numericamente las ecuaciones que involucren absolutamente todos los elementos que influyen sobre un evento, lo que significa la necesidad de suponer que los efectos de algunos elementos son despreciables, conduciendo a una aproximación.

Dado que una simulación involucra necesariamente aproximaciones, en nosotros está el poder de elegir cuales aproximaciones son las que simplifican el problema manteniendo las características esenciales. Hecho esto se transforma en un problema a resolver numericamente.

4.2 El papel de las simulaciones

“Hay muchos ejemplos de teorías viejas e incorrectas que persisten tercamente, mantenidas sólo por el prestigio de tontos pero bien conectados científicos... Muchas de estas teorías han sido muertas sólo cuando algún experimento decisivo expuso sus fallas... el trabajo del soldado en cualquier ciencia; en especial la física, es hecho por el experimental, quien mantiene honestos a los teóricos”, Michio Kaku [Dic99].

Tradicionalmente la física se ha caracterizado por una relación intensa entre experimentos y teoría, pero manteniendo siempre presentes importantes diferencias: En un experimento un sistema es sujeto de mediciones y los resultados son expresados comunmente en forma numérica. En la teoría se construye un modelo del sistema, usualmente (pero no siempre) en forma de ecuaciones matemáticas. Se verifica entonces la validez de este modelo según su capacidad para describir el comportamiento de un sistema en casos particulares en los que es posible calcular una solución de las ecuaciones y además existan experimentos con los cuales comparar. Una enorme cantidad de modelos interesantes (tanto en el ámbito académico como en el práctico) no eran verificables.

Con la aparición de las computadoras y su uso generalizado a partir de los años 50, se agregó un nuevo elemento, entre el experimento y la teoría: *la simulación*, también llamada *experimento por computadora*. En una simulación aún se propone un modelo, pero los cálculos los lleva a cabo la máquina siguiendo la “receta” del modelo (el *algoritmo* implementado en un *lenguaje de programación* en particular).

La aparición de la simulación altera la relación entre la teoría y el experimento en forma definitiva. Por un lado las simulaciones aumentan la demanda de precisión de los modelos. Por ejemplo, una simulación por dinámica molecular permite calcular la temperatura de fusión de un material con cierto modelo de interacción. Esta es una prueba complicada para el modelo “trayendolo a la vida” y permitiendo evidenciar sus áreas críticas (fuertes y débiles), dando lugar a sugerencias para mejorarlo.

Por otro lado, la simulación puede ser muy cercana al experimento, al considerar condiciones de tipo experimental, lo que permite comparar los resultados de la simulación con experimentos. La simulación también permite estudiar condiciones inaccesibles experimentalmente, ya sea por limitaciones físicas, económicas o de cualquier otro tipo; por ejemplo, sistemas a muy altas presiones o temperaturas, tiempos muy cortos o largos, escalas espaciales enormes o minúsculas, etc.

La simulación permite llevar a cabo “experimentos pensados” de situaciones imposibles en la realidad, cuyo resultado puede ayudar a entender mejor un fenómeno. Para hacer simulación es necesario traducir al lenguaje de programación una idea ajena a él en forma eficiente, pero también es necesaria una cierta dosis de creatividad y fantasía para sugerir experimentos y métodos.

Sin embargo, la simulación no es un sustituto de la experimentación, debido a que para simular hemos simplificado al sistema, eliminando factores cuyo efecto se desconoce y puede ser esencial. Nunca será posible eliminar la experimentación porque sólo a través de ella se observa al mundo “real” y este es infinitamente más complejo de lo que cualquier teoría puede considerar.

4.2.1 ¿Teoría o experimento?

“Inclusive si solo hubiera una posible teoría unificada, es sólo un conjunto de reglas y ecuaciones. ¿Que da vida a las ecuaciones y lugar a un universo para que describan? El acercamiento usual de la ciencia de construir un modelo matemático no puede responder las preguntas de ¿porque debe haber un universo para describir? ¿porque el universo se toma la molestia de existir?”, Stephen Hawking [Haw88].

Algunas veces se considera a la simulación como teoría, otras veces como experimento. Por un lado podemos decir que trata con modelos y no con la “realidad”, lo que sugiere clasificarla dentro de los métodos teóricos. Por otro lado, el procedimiento de verificar un modelo por medio de una simulación por computadora se asemeja más a un experimento: hacemos una corrida y luego analizamos los resultados en una forma muy parecida a como lo hacen los físicos experimentales. ¿De que lado clasificamos a la simulación?

No hay una respuesta definitiva a esta pregunta, pero no es necesaria. Ambos lados representan puntos de vista legítimos, esto es lo que hace de la ciencia computacional una rama interesante por si misma.

La teoría se basa, tradicionalmente, en un acercamiento *reduccionista*: tratamos con sistemas complejos reduciendolos a subsistemas simples, tanto como sea necesario para obtener sistemas suficientemente simples como para resolverlos. Cuando vemos a la simulación como una herramienta práctica para “verificar y probar” modelos en situaciones difíciles de tratar en forma analítica, estamos asumiendo que es el modelo el que se encuentra en el foco del interés.

Es importante darnos cuenta de que la simulación puede jugar un papel más importante e interesante. Podemos dejar de considerarla como un apoyo al sistema reduccionista, y más bien como una alternativa. La simulación puede ayudar a reducir la enorme diferencia entre los modelos que se pueden solucionar y los que no. Gracias a la simulación no tenemos que tratar con modelos tan simplificados como en el pasado.

Los potenciales intratómicos son un buen ejemplo. En el pasado las interacciones se suponían de pares con forma analítica simple, como los de Morse y Lennard-Jones. Hoy en día los más precisos contienen términos de muchos cuerpos y se determinan en forma numérica para reproducir, tanto como sea posible, fuerzas predichas por modelos de primeros principios. Nos hemos movido entonces un nivel arriba en el grado de reduccionismo contenido por el potencial. Estos nuevos potenciales no podrían existir sin la simulación, y esta deja de ser sólo una unión entre el experimento y la teoría para convertirse en una herramienta poderosa que abre nuevas opciones.

De ser necesario decidir si la simulación pertenece a la teoría o a la experimentación mi opinión personal es que pertenece a la teoría. La experimentación involucra una observación de la naturaleza, y en la simulación no

observamos la naturaleza sino su representación por ecuaciones fundamentales conseguidas por medio de la teoría. El análisis de datos experimentales inclusive de los más simples es ya parte de la teoría.

4.3 Diferentes métodos para simular

En la base de la simulación está el modelo físico del sistema. Usualmente nuestra intención es calcular algún promedio sobre un conjunto de muestras; por ejemplo, el promedio de la energía cinética que permite calcular la temperatura.

Si asumimos que el sistema tiene un hamiltoniano \mathcal{H} , denotamos al estado del sistema con $\mathbf{x} = (x_1, \dots, x_n)$, en donde n es el número de grados de libertad y al conjunto de estados accesibles en el espacio fase Ω . Una propiedad \mathcal{A} será una función de los estados del sistema, y lo que necesitamos especificar para poder calcularla es la función de distribución $f(\mathbf{x})$ del ensamble apropiado. \mathcal{A} queda dada entonces por

$$\langle \mathcal{A} \rangle = Z^{-1} \int_{\Omega} \mathcal{A}(\mathbf{x}) f(\mathcal{H}(\mathbf{x})) d\mathbf{x} \quad (4.1)$$

en donde

$$Z = \int_{\Omega} f(\mathcal{H}(\mathbf{x})) d\mathbf{x}$$

y se conoce como el promedio del ensamble, con la función de partición Z .

En algunos casos el promedio del ensamble se calcula propagándose al sistema a lo largo de una trayectoria en el espacio fase, evaluándose una propiedad \mathcal{A} mediante el promedio temporal

$$\bar{\mathcal{A}}_t = (t - t_0)^{-1} \int_{t_0}^t \mathcal{A}(\mathbf{x}(\tau)) d\tau \quad (4.2)$$

En otros se evalúa para un gran número de sistemas similares. Surge entonces una interesante pregunta: ¿son iguales ambos promedios?. La respuesta *no se sabe*, aunque la evidencia a favor de una respuesta positiva es abrumadora. Cómo no ha sido demostrado con absoluta rigurosidad aún mantiene el carácter de hipótesis, y es conocida con el nombre de *hipótesis ergódica*, que dice

$$\bar{\mathcal{A}}_{\infty} = \langle \mathcal{A} \rangle \quad (4.3)$$

Aquí surge una de las dos limitantes importantes de las simulaciones por computadora: no es posible seguir una trayectoria durante un tiempo infinito, así que el tiempo de observación está limitado a ser finito, de tal manera que el espacio fase sólo es muestreado, y uno debe contentarse con tener

$$\bar{\mathcal{A}} \simeq \langle \mathcal{A} \rangle \quad (4.4)$$

Hay algunos problemas para los que el tiempo finito de observación puede considerarse infinito (ver sección 4.6.3).

Si en una simulación se propaga al sistema en el espacio fase y de la forma de propagación depende como se calculan las propiedades, cabe hacernos la pregunta ¿cómo se propaga?. Para hacerlo distinguimos dos métodos:

1. Métodos deterministas.
2. Métodos estocásticos.

La idea detrás de los métodos deterministas consiste en utilizar la dinámica intrínseca al modelo para propagarlo. Se proponen las ecuaciones de movimiento para un conjunto de partículas y sus condiciones iniciales, y se integran en el tiempo, obteniendo una trayectoria en el espacio fase.

Los métodos estocásticos parten de suponer que lo único que se necesita evaluar son las posiciones, que la información referente al momento se puede obtener integrando. El problema entonces es cómo inducir las transiciones de una configuración a otra. Las transiciones se construyen según una evolución probabilista vía un proceso de Markov. El proceso de Markov es el análogo probabilista de las dinámicas intrínsecas, pero tiene la ventaja de permitir tratar sistemas que no tienen una dinámica intrínseca asociada.

Florian Müller-Plathe [Mül97] [Mül98] presenta una clasificación particularmente clara de los diferentes métodos de simulación molecular:

Supongamos que tenemos la función de energía potencial $\mathcal{V}(\mathbf{r})$. Su forma analítica simple nos permite evaluar sus gradiente con respecto a la posición de cualquier átomo i

$$\nabla_i \mathcal{V}(\mathbf{r}) = \frac{\partial \mathcal{V}(\mathbf{r})}{\partial \mathbf{r}_i} \quad (4.5)$$

Diferentes métodos de simulación molecular usan de diferentes formas la energía potencial y/o su gradiente, cuyo negativo es la fuerza. La figura 4-1 es un diagrama para ilustrar la siguiente clasificación.

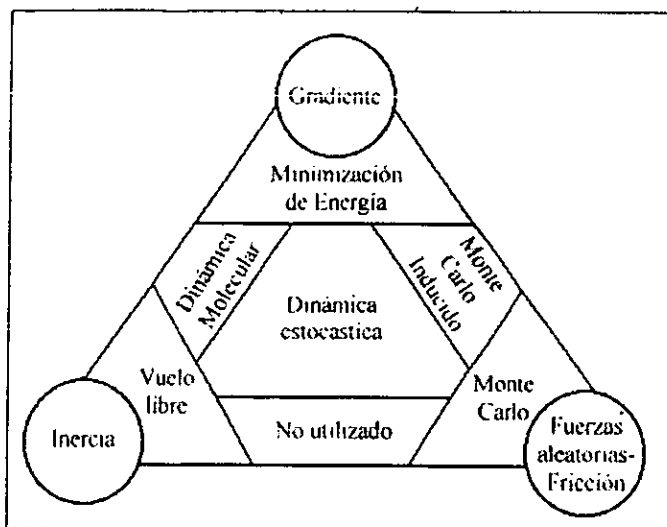


Figura 4-1 Diagrama de la relación entre diferentes métodos de simulación.

Si sólo se usa la información del gradiente, el sistema es llevado hacia el mínimo local más cercano. Esto es llamado *minimización de energía*, optimización geométrica. Algunos métodos usan derivadas de mayor orden, como el Hessiano, para determinar la calidad de los puntos de equilibrio encontrados y determinar si son estables (mínimos) o inestables (puntos de inflexión).

Si los átomos tienen alguna sensación de sus masas además del gradiente de la energía potencial, su movimiento estará en parte determinado por la inercia. En el caso más simple los átomos siguen la segunda ley de Newton, que en coordenadas cartesianas dice:

$$\mathbf{f}_i = m_i \mathbf{a}_i, \text{ o más formalmente, } \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = -\frac{1}{m_i} \nabla_i \mathcal{V}(\mathbf{r}_i). \quad (4.6)$$

Esta ecuación de movimiento lleva al método de *dinámica molecular*.

Si los átomos siguen las ecuaciones de movimiento de Newton, pero sin sentir ninguna fuerza, siguen lo que se llama vuelo libre. Para un ensamble de átomos este sistema se conoce como gas ideal, sin embargo es de reducida importancia en el contexto de simulación molecular, porque todo puede ser calculado en forma analítica.

El tercer componente es el de los movimientos aleatorios. Es posible mover átomos sobre una superficie de energía en una secuencia aleatoria y aún así obtener distribuciones que correspondan con ensambles estadísticos bien definidos. Este procedimiento no necesita del gradiente, basta con la energía potencial. Es conocido como *Monte Carlo* (o *Metropolis*). Durante una simulación Monte Carlo no se almacena información de las velocidades de los átomos, por lo que el sistema no tiene ninguna memoria de sus estados previos. Esto significa que Monte Carlo es un proceso de Markov. Esto puede ser visto también como fricción infinita: la velocidad inercial cae a cero y una fuerza aleatoria produce el siguiente estado. Sin embargo, Monte Carlo puede ser más eficiente si se utiliza información del gradiente para guiar los movimientos aleatorios. Este tipo de métodos se conoce como *Monte Carlo inducido por fuerza* (force-bias Monte Carlo).

Es posible idear ecuaciones de movimiento que contengan los tres ingredientes: gradiente, inercia y fuerzas aleatorias. Métodos de este tipo son la *dinámica browniana*, *dinámica de Langevin*, y se agrupan bajo *dinámica estocástica*. Una manera de verlos es pensar en átomos moviéndose en un líquido viscoso: Los átomos son acelerados por fuerzas sistemáticas (el gradiente), quieren ir en una misma dirección (inercia), pero sufren colisiones con los (supuestos) átomos del medio viscoso. El medio produce tanto fuerzas aleatorias que aceleran los átomos como una fricción que los desacelera. El cambio en las ecuaciones de movimiento puede ser

$$\mathbf{f}_i = m \mathbf{a}_i + \mathbf{x}_i + \gamma \mathbf{v}_i \quad (4.7)$$

donde x_i es una fuerza estocástica (aleatoria) y γv la fuerza de fricción. Sus magnitudes se acoplan de manera que la temperatura T del sistema se mantenga constante. El efecto es una transición del ensamble micro-canónico al ensamble canónico (ver sección 6.12).

4.4 Historia de la Dinámica Molecular

“La investigación es como caminar a ciegas en la oscuridad buscando el switch de la luz, una vez que es encendido todos pueden ver”, Desconocido [Ali99].

Un desarrollo completo de la historia de las técnicas de dinámica molecular sería demasiado largo y complicado para lo que pretendo en mi tesis, sin embargo es interesante incluir un recuento de los artículos principales aparecidos en los 50s y 60s que pueden ser considerados como las piezas angulares de la dinámica molecular. No incluyo artículos del desarrollo de los Métodos de Monte Carlo que ocurrieron en el mismo período (El lector interesado puede encontrarlos en [Woo57]).

Reimpresiones de todos estos artículos y de otros trabajos importantes en el área de la simulación por computadora de líquidos y sólidos hasta 1986 pueden encontrarse en [Cic87].

- El primer artículo reportando una simulación por dinámica molecular fué escrito por Alder y Wainwright [Ald57] en 1957. El propósito del artículo era investigar el diagrama de fase de un sistema de esferas duras, en particular de las regiones de sólido y líquido. En un sistema de esferas duras las partículas interactúan vía colisiones instantáneas y viajan como partículas libres entre colisiones (sección 7.4.2). Los cálculos los realizaron en una UNIVAC y una IBM 704.
- El artículo *Dynamics of radiation damage* (dinámica del daño por radiaciones) de J. B. Gibson, A.N. Goland, M.Milgram y G.H. Vineyard del Laboratorio Nacional en Brookhaven apareció en 1960 [Gib60] y es probablemente el primer ejemplo de cálculos de dinámica molecular con un potencial continuo de Born-Mayer de pares (sección 7.4): $\varphi = Be^{-\beta r}$, basado en un método de integración de diferencias finitas. El cálculo era en un sistema de 500 átomos, que en una IBM 704 requería de aproximadamente 1 minuto por paso. El artículo trataba de la creación de defectos inducidos por radiación (un tema de mucho interés y grandes presupuestos en los años de la guerra fría), es muy bueno, completo y útil por los resultados presentados. Sorprendente para tener 40 años!
- Anesseeur Rahman en el Laboratorio Nacional en Argonne es conocido como un pionero de la dinámica molecular. En su artículo de 1964 *Correlations in the motion of atoms in liquid argon* (correlación en el movimiento de átomos en argón líquido) [Rah64] estudia algunas propiedades del argón líquido bajo el potencial de Lennard-Jones en un sistema de 864 átomos usando una computadora CDC-3600. El legado del código de Rahman aún se usa en muchos programas[Erc97].
- Loup Verlet calculó en 1967 [Ver67] el diagrama de fase del argón usando un potencial de Lennard-Jones (sección 7.4.3), además de funciones de correlación para probar teorías del estado líquido. El método que más tarde sería conocido como *lista de vecinos de Verlet* es introducido en este artículo. También usa el algoritmo de integración de Verlet Velocidad (sección 6.2.5) muy importante hoy en día. Transiciones de fase en el mismo sistema fueron estudiadas por Hansen y Verlet [Han69] algunos años después.

4.5 Dinámica Molecular actual

La lista no es completa y refleja los temas sobre los que he investigado recientemente (materia condensada).

Líquidos. Como se ve en la sección 4.4, es en los líquidos donde todo comenzó. Siguen siendo un tema importante. La posibilidad de usar modelos de interacción más complejos y realistas permite estudiar nuevos sistemas, tanto elementales como de varios componentes. Por medio de técnicas para sistemas fuera de equilibrio [Cic86] se pueden estudiar fenómenos de transporte como viscosidad y flujo de calor.

Defectos. Otro campo de estudio investigado por la dinámica molecular desde sus comienzos [Cic87] son los defectos en cristales, cruciales por sus propiedades mecánicas y por lo tanto por su importancia tecnológica. El énfasis ha pasado de los defectos puntuales (vacancias e intersticiales) a lineales (dislocaciones) y planos (fallas de apilamiento y por supuesto las fronteras de grano tratadas en esta tesis) [Yip]. Nuevamente una de las fuerzas que mantiene viva el área es la de la mejora en el realismo del tratamiento gracias a mejores potenciales y a la posibilidad de tratar con sistemas de mayor tamaño.

Fractura. Bajo acción mecánica los sólidos se pueden romper en dos o más piezas. El proceso de fractura puede ocurrir de diferentes formas y con diferentes velocidades. La importancia tecnológica es obvia y a pesar de lo complejo del problema, la simulación nos da alguna idea de lo que ocurre.

Superficies. La física de superficies sufrió una explosión en los 80 gracias a la aparición de nuevas herramientas experimentales con resolución atómica (microscopía por tunelaje, microscopía de alta resolución y diferentes técnicas de dispersión). La simulación aún juega un papel importante en entender fenómenos como reconstrucción de superficies, fusión superficial, difusión, endurecimiento y otras que normalmente requieren de muestras y tiempos de simulación grandes.

Fricción. Aún más recientes son las investigaciones de adhesión y fricción entre dos sólidos, impulsadas por el desarrollo de la microscopía de fuerza atómica.

Cúmulos. Un cúmulo (*cluster* en inglés) es un conglomerado, que puede estar compuesto desde unos cuantos hasta varios miles de átomos o moléculas. Constituye un puente entre sistemas moleculares y sólidos comprendiendo algunas propiedades interesantes y complicadas. Frecuentemente un número sorprendentemente grande de configuraciones diferentes pueden tener energías similares, haciendo muy difícil determinar las estructuras más estables. Sus propiedades de fusión pueden ser radicalmente diferentes de las del sólido debido a su tamaño, a la anisotropía y sobre todo al papel que juega la gran proporción de superficie (en comparación con el sólido que normalmente se considera sin superficie). Los cúmulos metálicos son extremadamente importantes desde el punto de vista tecnológico debido a su papel como catalizadores en reacciones químicas importantes (por ejemplo en el escape de autos).

Biomoléculas. La dinámica molecular permite estudiar un gran número de macromoléculas, incluyendo sistemas biológicos como proteínas [Mon94], ácidos nucleicos (ADN y ARN), membranas [Ber87]. Eventos dinámicos pueden jugar un papel importante en el control de procesos que afectan las propiedades funcionales de la biomolécula. También hay aplicaciones importantes para la industria farmacéutica, inclusive en el *diseño de medicinas*, probando propiedades de una molécula en la computadora bajo condiciones específicas sin necesidad de sintetizarla y experimentar, lo que (de ser posible) resulta muy caro.

Propiedades y dinámica electrónicas. El desarrollo del método de Car-Parinello en el que las fuerzas sobre los átomos se obtienen al solucionar el problema de la estructura electrónica en lugar del potencial interatómico. Esto permite estudiar propiedades electrónicas de materiales en forma más completa incluyendo su dinámica y con ello transiciones de fase y otros fenómenos dependientes de la temperatura. Este tipo de trabajos ha dado lugar a una línea de investigación muy exitosa en la última década. Ver las secciones 4.6.2 y 7.9.

4.6 Limitaciones de la Dinámica Molecular

“Solo dos cosas son infinitas, el universo y la estupidez humana, y no estoy seguro de la primera.” Albert Einstein.

A pesar de ser una herramienta muy poderosa, la dinámica molecular tiene limitaciones importantes, y hay que tenerlas en mente al hacer simulación y sobre todo al sacar conclusiones. A continuación presento algunas de las limitaciones clave, generales a diferentes tipos de dinámica molecular y que además se relacionan con esta tesis.

4.6.1 Sólo mecánica clásica

¿Cómo podemos usar las ecuaciones de Newton para mover átomos cuando hoy en día hay teorías como la cuántica y la relatividad que sabemos son más precisas que la newtoniana?

Tomar en cuenta a la relatividad lo podemos posponer porque suponemos que las velocidades involucradas son muy pequeñas, y las correcciones son despreciables con respecto a otros errores (ver sección 7.1). Si queremos tomar en cuenta a los electrones y sus movimientos puede ser necesario incluir efectos relativistas pues estos se pueden mover a muy altas velocidades, sin embargo si nos restringimos a los movimientos de los núcleos atómicos (sección

7.1) entonces la aproximación de las ecuaciones de movimiento de Newton ha mostrado ser suficientemente buena como para dar resultados correctos, siempre y cuando las distancias involucradas no impliquen efectos cuánticos importantes.

Al simular átomos y moléculas no siempre podremos despreciar los efectos cuánticos, porque sus efectos son mayores conforme menores son las distancias involucradas. Sabemos que a nivel atómico los sistemas no obedecen las leyes de Newton, y la ecuación de Schrödinger es hasta ahora lo mejor.

Una prueba simple de la validez de la aproximación clásica se basa en la longitud de onda térmica de de Broglie [Han86] definida como

$$\Lambda = \sqrt{\frac{2\pi\hbar^2}{Mk_B T}} \quad (4.8)$$

en donde M es la masa atómica y T la temperatura. La aproximación clásica se justifica si $\Lambda \ll a$ en donde a es la distancia promedio a primeros vecinos. Si uno considera, por ejemplo, líquidos en el punto triple, Λ/a es de orden 0.1 para elementos ligeros como el Litio y el Argón, decreciendo conforme el elemento se hace más pesado. La aproximación clásica es, entonces, muy mala para elementos muy ligeros como H_2 , He y Ne y mejor para elementos más pesados como Au y Ag .

Los efectos cuánticos se vuelven más importantes en un sistema conforme T es más baja. La caída en el calor específico de cristales por debajo de la temperatura de Debye [Kit86] o el comportamiento anómalo del coeficiente de expansión térmica son ejemplos conocidos de efectos cuánticos medibles en sólidos [Chr88]. Es importante interpretar cuidadosamente los resultados de dinámicas moleculares en estas regiones.

4.6.2 Realismo de los potenciales usados

Las interacciones entre los átomos dan origen a fuerzas instantáneas que junto con las condiciones iniciales dictan la evolución de la configuración del sistema. Al moverse los átomos sus posiciones relativas cambian y con ellas cambian también las fuerzas.

Cómo explico en la sección 5.2 las fuerzas se obtienen como el gradiente de la energía potencial, que es función de las posiciones de las partículas (y en ocasiones de sus velocidades). El ingrediente físico principal será entonces el potencial (ver sección 7). Podemos decir que una simulación será más realista (reproduciendo más fielmente el comportamiento de un sistema real) conforme las fuerzas atómicas representen mejor las fuerzas interatómicas reales.

4.6.3 Limitantes de tiempo

Simulaciones de dinámica molecular típicas pueden involucrar sistemas con miles (o incluso millones) de átomos, por tiempos de simulación que van desde algunos picosegundos ($\sim 10^{-12}$ s) hasta cientos de nanosegundos ($\sim 10^{-7}$ s).

Se acostumbra decir que una simulación es "segura" desde el punto de vista de duración cuando el tiempo de simulación es mucho mayor que el tiempo de relajación de las cantidades que nos interesan. Sin embargo, diferentes propiedades tienen diferentes tiempos de relajación. Pueden requerir más tiempo para las relajaciones cuando se encuentran cerca de transiciones de fase, y no es raro encontrar casos en que los tiempos de relajación de alguna propiedad física sean varios ordenes de magnitud mayor que lo accesible.

4.6.4 Limitantes de tamaño

Simular un sistema finito también puede constituir un problema. En este caso uno debe comparar el tamaño de la celda de dinámica molecular con las longitudes de correlación de las funciones de correlación espaciales que le interesan. Al igual que con los tiempos de relajación la longitud de estas correlaciones puede crecer e incluso diverger en las proximidades de las transiciones de fase y los resultados dejan de ser confiables cuando se vuelven comparables con la longitud de la celda.

El problema puede ser tratado en forma parcial por el método conocido como *escalamiento finito*. El método consiste en calcular una propiedad física \mathcal{A} usando diferentes celdas con diferentes tamaños L , y entonces ajustar los resultados con la relación

$$A(L) = A_0 + \frac{c}{L^n} \quad (4.9)$$

usando A_0 , c y n como parámetros de ajuste. A_0 corresponde entonces al límite $\lim_{L \rightarrow \infty} A(L)$, y debe ser tomado como un estimable bueno de la cantidad física "real".

Esta relación permite buscar una convergencia hacia un valor físico "real", pero no necesariamente nos ha de entregar resultados correctos (desde el punto de vista de su congruencia con la realidad) debido a fenómenos colectivos no importantes en sistemas pequeños y que pueden ser de importancia definitiva conforme se aumenta el tamaño de la muestra. Afortunadamente existen otros métodos para tratar con este problema, los más comunes son las condiciones periódicas (ver sección 6.3) en donde se busca reproducir un sistema infinito. Debido a que la celda simulada es en efecto finita, estos métodos logran evitar algunos de los problemas de reducir el sistema en tamaño, pero no todos, y aún encontramos problemas al tratar con longitudes de correlación mayores que las dimensiones de la celda. Hay que tener cuidado de no intentar obtener resultados de correlaciones mayores que estas longitudes porque podemos encontrar efectos derivados de la imposición que hicimos de periodicidad y no de propiedades del sistema.

En la sección 6.3 trato con este problema en forma más detallada, presentando soluciones por simetría traslacional (sección 6.3.4), condiciones de frontera esféricas (sección 6.3.6) y la forma de calcular las fuerzas al aplicar estas condiciones (sección 6.3.7).

El tamaño finito del sistema puede tener ventajas, por ejemplo para calcular la energía libre como muestro en la sección 6.6.12.

4.6.5 Limitaciones en precisión al reproducir sistemas dinámicos

Hace unos 30 años Lorenz descubre lo que ahora llamamos Teoría del caos. Uno de los conceptos interesantes de la teoría del caos, es el de Sensibilidad a condiciones iniciales [Ban92] [Ben92]: si preparamos dos sistemas con condiciones iniciales ligeramente diferentes entre (en donde ligeramente significa que podemos hacerlas tan parecidas como queramos siempre y cuando sean diferentes), tarde o temprano al evolucionar los sistemas esta diferencia crecerá. ¿Que significa esto?, significa que aún cuando la interpretación estándar de la mecánica cuántica no tenga razón [Gas95] y tenga sentido hacer una teoría que describa con precisión absoluta las condiciones iniciales de un sistema, tendremos que por pequeño que sea nuestro error al medir alguna (cualquiera) de las variables, tarde o temprano al evolucionar el sistema el error habrá crecido y opacado los cálculos útiles y no habrá razones para afirmar que el sistema encontrado tenga algún parecido con el sistema real.

Cualquiera de los inevitables errores de precisión introducidos en los cálculos lleva a errores en los resultados y a posibles interpretaciones equivocadas. No se conocen los efectos dinámicos caóticos asociados a los sistemas tratados, pero por lo general se suponen despreciables. Si nuestra intención es únicamente la de hacer un muestreo del espacio fase para calcular propiedades del sistema, entonces la sensibilidad a condiciones iniciales no acarrea problemas, siempre y cuando tengamos en mente que la evolución temporal del sistema simulado no es necesariamente fiel a la realidad descrita por las ecuaciones fundamentales.

4.7 Monte Carlo

Aunque esta tesis centra en los métodos de dinámica molecular, incluyo las ideas básicas de los algoritmos de Monte Carlo porque es importante conocer un método alternativo y porque es necesario entender la forma en que se hacen simulaciones con cuyos resultados habremos de comparar los nuestros.

El método de Monte Carlo fué propuesto por Metropolis [Met49] por su extenso uso de números aleatorios, y ha sido utilizado en simulaciones por computadora desde los comienzos de los años 50. En este método se genera una configuración inicial (posiciones solamente) (sección 5.1). Se calcula la energía, y comienza un proceso repetitivo, cuyos pasos son [Car96]:

1. Generar un pequeño cambio en la configuración. En la simulación clásica de Metropolis [Met53], este era un desplazamiento elegido al azar de una partícula también elegida al azar. En otras variantes de Monte Carlo el desplazamiento es en la dirección de la fuerza sobre la partícula.
2. Calcular la energía de la nueva configuración.
3. Aceptar la nueva configuración si tiene una menor energía que la anterior, de lo contrario, aceptarla con una probabilidad

$$e^{-\frac{\mathcal{E}_{nueva} - \mathcal{E}_{vieja}}{k_B T}} \quad (4.10)$$

donde \mathcal{E}_{vieja} y \mathcal{E}_{nueva} son las energías de las configuraciones vieja y nueva respectivamente, k_B es la constante de Boltzmann y T es la temperatura absoluta. Una modificación natural de esto sería aceptar las configuraciones de menor energía bajo el mismo principio.

4. Calcular las propiedades deseadas, recolectar datos para promediar, guardar configuraciones, etc.

Los pasos se repiten tantas veces como sea necesario. Se puede notar que la nueva configuración depende directamente sólo de la configuración anterior, por lo que la secuencia de configuraciones forma una cadena de Markov. El esquema Monte Carlo permite medir propiedades de equilibrio, pero no muestra la dinámica que lleva a él. No hay información inherente que contenga el tiempo de simulación, aunque comparaciones con experimentos pueden dar estimaciones. Se acostumbra suponer que ha transcurrido un intervalo dt cuando en promedio cada átomo ha sido movido una vez.

La clave para que el método funcione es un buen generador de números aleatorios, por lo que la mayoría de los libros que tratan el tema incluyen un apéndice al respecto [All87] [Hee90].

5. LA MAQUINARIA BÁSICA DE LA DINÁMICA MOLECULAR

“El uso de la fuerza bruta es tan mala solución de los problemas que generalmente sólo lo usan los niños chiquitos y las grandes potencias.” David Friedman

El método de simulación de dinámica molecular consiste en lo siguiente:

1. Generar un sistema inicial, que en el caso de átomos consiste de las posiciones y velocidades iniciales de cada uno de los átomos, de tal manera que sean consistentes con un estado inicial determinado calculable según sus propiedades termodinámicas,
2. Suponer un potencial de interacción entre los elementos. Junto con las leyes de movimiento permite a la simulación reproducir sistemas “reales”.
3. Integrar las ecuaciones resultado de las leyes de movimiento y el potencial de interacción para observar la evolución del sistema.

A continuación describo brevemente cada uno de los puntos, mostrando algunas de sus propiedades principales y de la forma en la que nos aproximamos a ellos para hacer dinámica molecular.

Durante la integración se realizan diferentes mediciones de propiedades del sistema, que permiten modificarlo y “forzar” la evolución en la dirección que nos interese.

5.1 Configuración

“Los cristales son como las personas: son sus defectos los que los hacen interesantes” Carlos Ruiz.

Una configuración es un estado del sistema, determinado tanto por las posiciones como velocidades atómicas. De la configuración inicial partimos para integrar las ecuaciones de movimiento. Conocerla en forma exacta es, por lo general, imposible. Recordemos que configuraciones iniciales muy parecidas pueden diverger con el tiempo (dejando ver propiedades caóticas de los algoritmos de integración) o converger (mostrando atractores) (secciones 4.6.5 y 6.2).

Hay diferentes maneras de especificar las posiciones y velocidades iniciales, la elección dependerá de los objetivos y de la forma en que se implementen. A continuación explico algunos de los procedimientos comunes.

5.1.1 Posiciones

Una de las primeras opciones que viene a la mente respecto a posiciones iniciales es hacerlas al azar. Es factible para casos de densidades bajas, pero al aumentar la densidad deja de ser una opción recomendable porque aumenta la probabilidad de tener átomos muy cercanos, entre los que las fuerzas sean de gran magnitud (inclusive infinitas para potenciales como el de esferas duras (sección 7.4.2)). lo que lleva a errores en los algoritmos de integración.

Cristales

Es común comenzar a simular utilizando potenciales simples (capítulo 7) con estructuras conocidas para elementos descritos por ellos. Un ejemplo bien conocido es la estructura FCC, por lo que es casi una convención validar un programa usándola.

Una red FCC se construye repitiendo un arreglo fundamental de átomos conocido como *celda unidad*. La celda unidad es una receta con la localización del menor número n de átomos, necesarios para definir la simetría de la

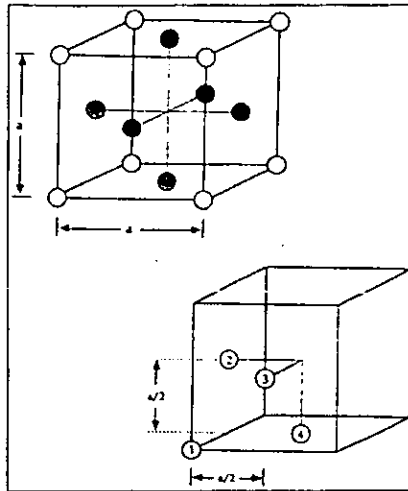


Figura 5-1 Arriba se muestra un cubo de una red FCC y lado a . En cada cara del cubo hay un átomo. Abajo los cuatro átomos que definen la celda unidad FCC. [Hai92]

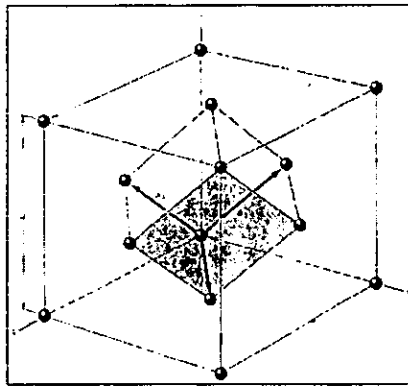


Figura 5-2 La celda primitiva y el romboedro que forma. a_1 , a_2 y a_3 son los vectores de traslación. [Kit86]

red [Hai92]. Para estructuras cúbicas cada lado de la celda unidad tiene longitud a y la densidad de la celda es la densidad de la red completa $\rho = \frac{n}{a^3}$.

La celda unidad de la estructura FCC contiene cuatro átomos en las posiciones

$$a_1 = (0, 0, 0) \tag{5.1a}$$

$$a_2 = (0, \frac{1}{2}a, \frac{1}{2}a) \tag{5.1b}$$

$$a_3 = (\frac{1}{2}a, 0, \frac{1}{2}a) \tag{5.1c}$$

$$a_4 = (\frac{1}{2}a, \frac{1}{2}a, 0) \tag{5.1d}$$

en donde el origen se encuentra en uno de los vértices de la celda (figura 5-1). Las posiciones de los átomos anteriores definen tres *vectores primitivos de traslación*. Una red FCC completa puede ser construida usando los vectores definidos por cada átomo trasladándolos a en cada dirección cartesiana. Una red FCC completa debe tener entonces N átomos, donde se debe cumplir $N = 4I^3$ con I entero (4 por los átomos de cada repetición de la celda unidad e I^3 por las repeticiones en cada dirección). De tal manera que configuraciones iniciales de una red FCC normalmente usan 32, 108, 256, 500, 864, 1372, etc... átomos. La figura 5-2 muestra la celda romboedral primitiva definida por los vectores primitivos a_2 , a_3 y a_4 y nos muestra porque conocemos a la estructura como FCC: porque coincide con un cubo con átomos en los vértices y en el centro de las caras: de allí a su nombre FCC, del inglés para *Face Centered Cubic*.

Las redes pertenecen al sistema cúbico, uno de los 7 diferentes existentes para teselaciones en tres dimensiones

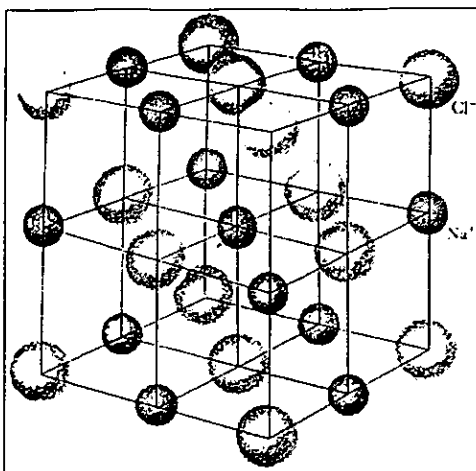


Figura 5-3 Representación de la estructura del Cloruro de sodio, alternando iones de sodio Na^+ y cloro Cl^- . Los diámetros iónicos están reducidos para hacer más claro el arreglo. [Kit86]

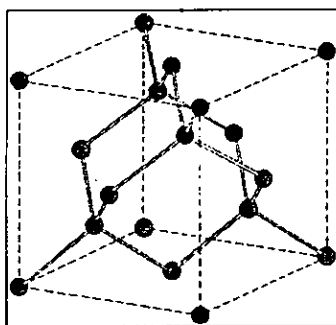


Figura 5-4 Estructura del cristal de diamante mostrando el arreglo tetraedral. [Kit86]

[Kit86]. Propiedad importante de una red cúbica son sus simetrías, por ejemplo: traslaciones respecto a cualquiera de los vectores primitivos dejan la red indistinguible, mostrando una simetría.

Hay muchos elementos que cristalizan en estructuras FCC, pero esta estructura no se limita a cristales monoatómicos, un ejemplo es el NaCl, en donde los átomos de sodio y cloro forman estructuras FCC monoatómicas interpenetradas, como se muestra en la figura 5-3. El diamante también forma una estructura basada en la FCC, pero con átomos extra en la posición $a_5 = (\frac{1}{4}a, \frac{1}{4}a, \frac{1}{4}a)$, lo que permite a los átomos de carbono unirse en forma covalente con la simetría adecuada a sus propiedades electrónicas. La estructura resultante se muestra en la figura 5-4. El átomo con vector primitivo a_5 puede ser sustituido por un 2° elemento para formar compuestos binarios.

Por el tipo de simetrías de una red FCC se clasifica junto con las redes Cúbica simple y Hexagonal. Para cualquiera de las ellas puntos de la red en planos perpendiculares a una diagonal del cubo (un eje de simetría 3) forman una red hexagonal bidimensional. De esta forma podemos explicar un método alternativo para construir una red FCC. El método está explicado en la figura 5-5 y explica también porque las redes FCC y HCP (del inglés para *Hexagonal Close Packed*) se consideran hermanas y tienen el mismo máximo empaquetamiento para esferas duras (sección 6.6.9).

Cúbica centrada en la cara (FCC)

Cuando queremos generar una configuración FCC podemos seguir la regla cristalográfica de construirla con la celda base repetida espacialmente con la simetría cúbica. De esta forma se acostumbra construir los clusters de átomos para simular. Dado que la celda base FCC tiene 4 átomos, es común que por razones de simetría se acostumbre generar configuraciones de 4, 32, 108, 256 átomos ($4 \frac{\text{átomos}}{\text{celda}} \times \# \text{ celdas}$). El número de celdas es el cubo de la cantidad de celdas base que tenga el cluster a generar.

El método es muy bueno y da lugar a sistemas que pueden ser simulados con condiciones de frontera periódicas. Sin embargo, cuando se quiere simular un cluster y no un cristal, puede ser más conveniente construir un sistema con

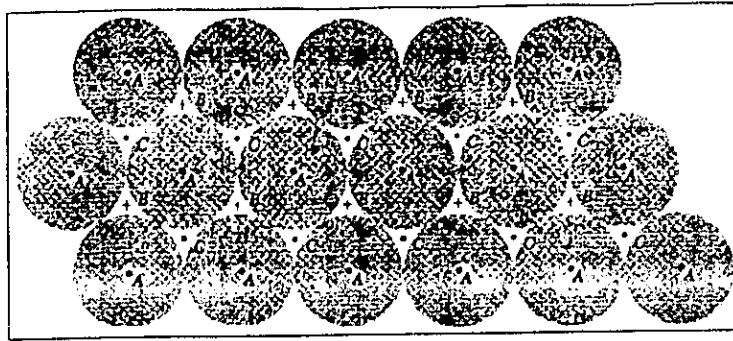


Figura 5-5 Los discos grises denotan una capa de esferas con centros sobre A. Se puede colocar arriba de estas una segunda capa idéntica y paralela al plano del dibujo, con centros sobre B. Hay dos opciones para una tercera capa: puede colocarse sobre A o sobre C. Si se hace sobre A la secuencia es ABABAB... y la estructura es HCP (Hexagonal Close-Packed). Si se coloca sobre C la secuencia es ABCABC... y la estructura es FCC (Face Centered-Cubic). [Kit86]

Orden	Coordenadas	Cantidad	Distancia al origen
1	$\pm\frac{1}{2}, \pm\frac{1}{2}, 0$	12	$\frac{1}{2}$
2	$\pm 1, 0, 0$	6	1
3	$\pm 1, \pm\frac{1}{2}, \pm\frac{1}{2}$	24	1.5
4	$\pm 1, \pm 1, 0$	12	2
5	$\pm\frac{3}{2}, \pm\frac{1}{2}, 0$	24	2.5
6	$\pm 1, \pm 1, \pm 1$	8	3
7	$\pm\frac{3}{2}, \pm\frac{3}{2}, \pm\frac{1}{2}$	48	3.5
8	$\pm 2, 0, 0$	6	4
9	$\pm\frac{5}{2}, \pm\frac{3}{2}, 0$	36	4.5
10	$\pm 2, \pm 2, \pm 1$	24	5

Tabla 5.1 Ejemplos de combinaciones de coordenadas de vecinos en una configuración FCC

un átomo en el origen y los demás colocados en sus posiciones cristalinias alrededor de este. De esta forma en lugar de generarse una configuración con simetría cúbica se da lugar a lo que llamamos una *cebolla*. El nombre viene de que después de colocar el átomo del origen vamos colocando los demás átomos llenando las capas de vecinos alrededor de este.

Un cristal FCC tiene 12 primeros vecinos, 6 segundos vecinos, 24 terceros vecinos, y así sucesivamente. En este método de construcción las cantidades comunes de átomos ausar son 13, 19, 43, 55, 79, etc. que son la suma de vecinos más el átomo central.

No es necesario hacer simulaciones con capas de vecinos llenas, pero igual que cuando se construyen repitiendo la celda base, por razones de simetría acostumbra hacerse.

La ventaja de crear un sistema de esta forma es que tenemos un cluster con la mayor cantidad posible de simetrías en un FCC para el átomo del origen.

Utilizando coordenadas tales que el parámetro de red sea 1 definimos *vecino de orden N* de un átomo en el origen de la siguiente manera:

1. Las coordenadas (x, y, z) de un átomo cumplen $x + y + z = \text{entero}$.
2. Definimos como vecinos de orden N a los átomos que se encuentran a una distancia $\sqrt{\frac{1}{2}N}$. Por lo tanto las coordenadas de un vecino de orden N cumplen $x^2 + y^2 + z^2 = \frac{N}{2}$.

La primera regla es ampliamente conocida en cristalografía, y es equivalente a afirmar que $x + y + z = \text{par}$ cuando el parámetro de red vale $\sqrt{2}$.

La segunda regla tiene como consecuencia que: El número de vecinos de orden N esta dado por las combinaciones de las coordenadas x, y y z que cumplan 1 y 2. Hasta donde conocemos la regla 2 no ha sido descrita en la bibliografía, y sin embargo permite la construcción de clusters con átomos en posiciones de un cristal FCC.

Como ejemplos de combinaciones de coordenadas esta la tabla 5.1.

Otra consecuencia interesante de la definición es que existen órdenes de vecinos de los que hay cero átomos. Este es un resultado importante, y damos un ejemplo: No existen valores x, y y z que satisfagan las reglas 1 y 2 anteriores

cuando $N = 14$. Igualmente encontramos que no existen valores que satisfagan $N = 30, 46, 62, 78, \dots$. Surge entonces la pregunta ¿cómo podemos saber para cuales órdenes hay cero vecinos?. La respuesta viene de la teoría de números:

No hay vecinos de orden N si existen j y k enteros tales que $2N = 4^j(8k + 7)$.

Dado que buscamos un resultado par sabemos que $j > 0$.

La definición usual de vecino no permite tal caso, sin embargo haciendo esta concesión la distancia a los vecinos de orden N es tan simple como $\sqrt{\frac{1}{2}N}$. Con la definición usual no conocemos ninguna regla simple que nos de la misma información. Según la definición usual los primeros 13 órdenes de vecinos si cumplen con la regla 2, pero los vecinos de orden 14 son 48 y están a $\sqrt{7.5}$ del origen. En nuestra definición de orden 14 hay 0 vecinos, y de orden 15 hay 48 a una distancia $\sqrt{7.5} = \sqrt{\frac{1}{2}15}$.

Si consideramos un cluster de 30,000 átomos encontramos que los vecinos con cero átomos pertenecientes a este cluster son los órdenes 14, 30, 46, 62, 78, 94, 110, 120, 126, 142, 206, 222, 224, 476, 248, 254, 270 y 286. Así el átomo 30,000 es, según nuestra definición, de orden 294 y según la definición usual de orden 276.

La figura 5-6 muestra la distribución de pares según esta definición. En la figura están incluidos los órdenes de vecinos para los que hay cero átomos, claramente visibles sobre el eje x.

Respecto a la cantidad de átomos de cada orden quisiera comentar que no es fácil predecirla para cualquier orden. La figura 5-6 muestra la gran variación de estos valores, y aunque no puede ser cualquier cantidad dadas las simetrías en la configuración, no es fácil anticiparla. Para los primeros órdenes es fácil calcular el número de vecinos porque son simples combinaciones de las coordenadas posibles, por ejemplo las de la tabla 5.1, pero para órdenes mayores vemos que no sólo son las combinaciones de un mismo tipo de coordenadas, sino que la misma distancia puede ser producto de combinaciones de coordenadas muy diferentes. Esto es interesante y valdría la pena buscar en el futuro alguna manera de predecir esta circunstancia.

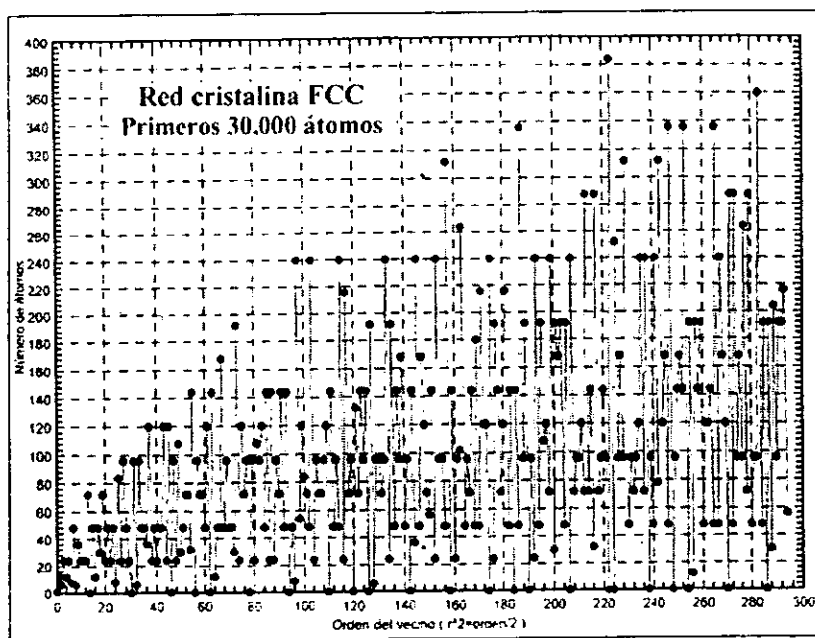


Figura 5-6 Distribución de pares para 30,000 átomos en FCC. El eje x no lo representa la distancia sino el orden del vecino con $orden = 2r^2$.

Cúmulos

Un cúmulo (*cluster* en inglés) es un conjunto. puede ser de átomos, moléculas, estrellas u otra cosa. Sabemos que las propiedades de un material pueden depender de su tamaño, y aún cuando la teoría de sólidos cristalinos está muy desarrollada, no es claro lo que ocurre cuando tratamos con sistemas pequeños. Es común que conjuntos pequeños tengan propiedades radicalmente distintas de las de conjuntos grandes. Conforme un cúmulo aumenta en tamaño esperamos que estas se diferencias desaparezcan (sección 6.3).

Líquidos

El método más simple de construir una estructura líquida es colocar átomos o moléculas en posiciones al azar dentro de una celda de simulación a repetir periódicamente (sección 6.3). Dado que una configuración construida de esta forma puede contener átomos demasiado cerca, representando situaciones poco físicas que llevan a problemas numéricos, en dinámica molecular es más común comenzar de una red y aumentar su energía cinética hasta llevarla más allá del punto de fusión (sección 6.3). Cualquier red es útil, pero históricamente las estructuras FCC han sido las más usadas. La escala de la red se elige de tal forma que se obtenga la densidad líquida deseada. (sección 6.6.9). Durante el curso de la simulación la estructura de red desaparece y se obtiene una estructura líquida típica. Para estimular el proceso de fusión se le puede dar a cada una de las partículas un pequeño desplazamiento fuera de su posición en la red.

Para el caso de fluidos moleculares es necesario especificar también la orientación de las moléculas.

Simetrizando

Una vez dadas las posiciones de los átomos es conveniente asegurarnos de tener un origen de coordenadas cómodo.

Una solución simple y efectiva, es utilizar el centro de masa \mathbf{R} definido como

$$\mathbf{R} = \frac{\sum_i^N m_i \mathbf{r}_i}{\sum_i^N m_i} = \frac{\sum_i^N m_i \mathbf{r}_i}{M} \quad (5.2)$$

donde he sustituido la masa total $\sum_i^N m_i$ por M . \mathbf{R} es particularmente simple cuando todos los átomos son idénticos y se ha cambiado la unidad de masa a la masa reducida 1:

$$\mathbf{R} = \frac{1}{N} \sum_i^N \mathbf{r}_i \quad (5.3)$$

De esta forma estamos sustituyendo el sistema de coordenadas por otro menos arbitrario, con respecto al cual podemos esperar conservación del momento.

Pequeñas desviaciones

Por la precisión finita de la computadora no siempre es posible asignar posiciones iniciales exactas, por lo que acostumbramos querer conocer los efectos que pequeñas desviaciones tienen en los resultados.

Una posibilidad es agregar pequeñas variaciones de las posiciones y observar sus efectos. Si después de varias pruebas con diferentes variaciones encontramos que el sistema no resulta muy dependiente de estas variaciones es probable que podamos dormir tranquilos a pesar de los errores agregados por la máquina durante los cálculos; pero si el sistema resulta sensible a pequeñas diferencias iniciales entonces no hemos de pegar el ojo hasta conocer el alcance e implicaciones de estas dependencias. Nosotros decidimos cuando el sistema es *muy* dependiente y cuando no.

Para hacer estas pruebas lo común es realizar simulaciones idénticas salvo por una pequeña modificación al azar en las posiciones iniciales. Un 1% suele ser razonable, aunque en ocasiones podemos pensar en modificarlas más dependiendo de que tan robusto queremos probar al sistema.

Para modificar las posiciones podemos:

1. Escalar las posiciones según un factor ligeramente diferente de uno.
2. Agregar desplazamientos pequeños a cada uno de los elementos, ya sea idénticos en magnitud pero de diferente dirección para cada átomo o inclusive de diferentes magnitudes.

Después de estas variaciones puede ser necesario volver a simetrizar el sistema (llevar el centro de masa al origen).

Modificar las posiciones, inclusive en muy pequeñas proporciones puede modificar ampliamente la energía potencial del sistema, basta poner como ejemplo el potencial de Lennard-Jones (sección 7.4.3) cuya parte repulsiva varía como r^{-12} , por lo que una ligerísima variación en la distancia interatómica puede tener efectos espeluznantes en la energía. Por lo mismo se acostumbra mantener las modificaciones pequeñas.

5.1.2 Velocidades

Hay que ser cuidadoso al asignarlas, pues podríamos dar demasiada energía cinética a algún átomo y hacer que este desordene sus alrededores a pesar de que la energía cinética promedio del sistema se encuentre por debajo de la energía de fusión.

Es necesario también tener cuidado con asignar velocidades iniciales que den un momento lineal y angular inicial diferente de cero (ver secciones 6.6.4 y 6.6.5). Esto sólo contribuye con una deriva general del sistema, complicando notablemente el cálculo y análisis de propiedades sin aportar nada nuevo.

Cuando se simulan cristales y se usan condiciones de frontera diferentes de la libre (sección 6.3) el momento angular no se conserva, por lo que no es de suma importancia hacerlo cero al inicio. De cualquier manera se acostumbra calcularlo de vez en cuando como una observación más de la evolución del sistema.

Cero

Si no queremos tener un sistema con momentos diferentes de cero ¿porqué no hacer todas las velocidades iniciales cero?. Porque la energía cinética también define la temperatura: una energía cinética inicial cero define una temperatura inicial cero. Si la energía cinética no es una constante tampoco lo será la temperatura, pero tener variaciones en la temperatura puede ser un problema si el sistema se encuentra en un estado próximo a una transición de fase. donde una fluctuación puede llevar a alguna parte del sistema más allá de la transición y complicar el análisis de resultados.

Asignar velocidades iniciales cero puede ser muy útil para probar algoritmos y verificar su operación antes de utilizar situaciones más complicadas. Es útil también como raíz de procesos de equilibrio, de esta forma: se parte de un sistema con velocidades cero, se integran las ecuaciones de movimiento por algunos intervalos, y (salvo que el sistema se encontrara al comienzo en un excelente mínimo) habrá ganado alguna energía cinética a cambio de una baja en la energía potencial. Modificamos entonces las velocidades para ir ajustando la energía cinética (y la temperatura) según lo descrito en la sección 6.12.

Un error común al pretender utilizar velocidades nulas es no asignarlas, en lugar de asignar velocidades cero. No hay que olvidar que no siempre al generar el espacio para almacenar las velocidades en la memoria de la computadora estos son inicializados en cero, y que pueden contener valores extraños y sin sentido. Es fácil, rápido y mucho más seguro asignar valores iniciales, aún siendo cero.

Proporcionales a la posición

Un poco más complejo que velocidades cero, y que genere un sistema con una cierta temperatura o energía cinética inicial diferente de cero, es hacer las velocidades proporcionales a la posición atómica. Digamos

$$\mathbf{v}_i = k\mathbf{r}_i \quad (5.4)$$

con k =constante. Podría ser que k fuera una función, por ejemplo decreciente de la distancia al origen, de tal manera que los átomos lejanos tengan velocidades pequeñas y los átomos cercanos velocidades grandes.

Este método es útil sobre todo en sistemas simétricos centrados en el origen, porque garantiza momentos iniciales cero.

Idénticas en magnitud

Otra posibilidad simple sería asignar velocidades idénticas en magnitud a todos los átomos. Si la distribución de direcciones es la adecuada, se puede conseguir que los momentos totales sean cero aún para un sistema que tenga energía o temperatura iniciales arbitrarias.

Este caso puede ser una generalización del caso anterior, con

$$\mathbf{v}_i = k_i\mathbf{r}_i \quad (5.5)$$

Con una distribución adecuada para la k_i . Esta distribución puede tener muchas formas, pero no todas son igualmente útiles. En la sección 6.4 explico que en un sistema en equilibrio debe ser posible partir el sistema arbitrariamente en dos y ambas mitades deben conservar propiedades del todo, esto pone limitaciones sobre la forma de k_i . Una forma simple para k_i podría ser una variación de la utilizada para poner velocidades proporcionales a la posición: $k_i = \frac{a}{r_i}$, de esta forma todos los átomos tienen la misma velocidad a .

Es útil utilizar una función k_i con dirección al azar y que asigne velocidades iguales, de tal manera que para sistemas grandes aún obtengamos momentos iniciales cero, pero no hayamos inducido al sistema alguna simetría por usar un origen arbitrario.

Distribución de Maxwell-Boltzmann

Una variación más es utilizar una función de azar para asignar tanto las direcciones como las magnitudes de las velocidades, obteniendo una cierta distribución inicial:

$$\rho(v_{ix}) = \left(\frac{m}{2\pi k_B T} \right)^{1/2} e^{-\frac{1}{2} m v_{ix}^2 / k_B T} \quad (5.6)$$

en donde $\rho(v_{ix})$ es la densidad de probabilidad para la componente x de la velocidad del átomo i , habiendo ecuaciones equivalentes para y y z . Es más complicado, pero puede ayudar a llevar el sistema más rápidamente al equilibrio, evitando peligrosas fluctuaciones.

Asignando un momento lineal inicial cero

Es recomendable que el momento lineal total inicial sea cero, para tener esto lo más común es hacer

$$v_{ix}^{nueva}(t=0) = v_{ix}^{actual}(t=0) - \frac{1}{N} \sum_j v_{jx}(t=0) \quad (5.7)$$

Así distribuimos la carga de la corrección en forma homogénea sobre todos los átomos. Esto modifica la magnitud de las velocidades y con ello la temperatura asignada originalmente, pero si el momento inicial era pequeño la modificación a la temperatura lo será también. Esta variación en la temperatura inicial puede ser corregida durante la fase de equilibrio (sección 6.4).

Pequeñas desviaciones

Para las velocidades no es común aplicar pequeñas desviaciones de las velocidades iniciales como hicimos con las posiciones (sección 5.1.1) ya que la mayoría de los métodos usados para asignar las velocidades involucran efectos de azar. Si nos interesara, el procedimiento sería equivalente al de las posiciones. Al igual que las variaciones de posición modificaron la energía potencial del sistema las variaciones de velocidades modificarán la energía cinética y la temperatura.

Asignando una energía cinética o temperatura inicial

Una vez asignadas las velocidades pueden ser escaladas para obtener una energía inicial objetivo $\mathcal{E}_{objetivo}$ mediante

$$v_{ix}^{nueva}(0) = v_{ix}^{vieja}(0) \sqrt{\frac{\mathcal{K}_{objetivo}}{\mathcal{K}_{actual}}} = v_{ix}^{vieja}(0) \sqrt{\frac{\mathcal{E}_{objetivo} - \mathcal{V}}{\mathcal{E}_{actual} - \mathcal{V}}} \quad (5.8)$$

en donde \mathcal{K} es la energía cinética y \mathcal{E} la total. Las componentes y y z se escalan en forma análoga.

En lugar de 5.8 podemos usar

$$v_{ix}^{nueva}(0) = v_{ix}^{vieja}(0) \sqrt{\frac{T_{objetivo}}{T_{actual}}} \quad (5.9)$$

para obtener una cierta temperatura inicial objetivo $T_{objetivo}$.

Al respecto del escalamiento de velocidades escribo más en la sección referente al control de la simulación y el sistema en la sección 6.12.

5.2 El potencial

"Mientras las leyes de las matemáticas se refieren a la realidad, no serán ciertas. Mientras sean ciertas no se referirán a la realidad." A. Einstein

El ingrediente principal de una simulación es el modelo del sistema físico. Para una simulación de dinámica molecular esto significa elegir un *potencial*. Un potencial es una función

$$\mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N), \quad (5.10)$$

de las posiciones y los momentos de los elementos del sistema, que representa un tipo de energía de esa configuración. Acostumbramos usar versiones simplificadas del potencial en que suponemos no hay dependencia de los momentos de las partículas. La función del potencial puede ser invariante ante traslaciones y rotaciones, en cuyo caso se acostumbra construir el potencial como una función de las posiciones relativas entre los átomos. Es posible agregar asimetrías para representar interacciones del sistema con sus alrededores.

Si el potencial depende únicamente de las posiciones relativas entre los átomos, las fuerzas son derivables como los gradientes del potencial con respecto a desplazamientos atómicos

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} \mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (5.11)$$

Esta forma implica la presencia de una ley de conservación de la energía total $\mathcal{E} = \mathcal{K} + \mathcal{V}$, donde \mathcal{K} es la energía cinética instantánea.

La forma más simple para elegir a \mathcal{V} es escribirlo como una suma de interacciones de pares

$$\mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_i^N \sum_{j>i}^N \phi(|\mathbf{r}_i - \mathbf{r}_j|) \quad (5.12)$$

en donde las sumas sobre $j > i$ e i significan que sólo ha de considerarse cada par de elementos una vez. En el pasado la mayoría de los potenciales eran interacciones de pares, pero hoy en día ya no es así.

Si el potencial es el ingrediente físico principal que permite una correlación con el mundo y una comparación con los experimentos, no es sorprendente que simplificaciones de esta tipo presenten muchas dificultades al predecir sistemas de interés práctico, por lo que los potenciales de pares han sido complementados y sustituidos de muchas formas. El desarrollo de potenciales adecuados representa una muy importante línea de investigación y resulta clave en las perspectivas de esta tesis. por lo que abundo en el en la sección 7.

5.3 Ecuaciones de movimiento

“De lo que voy a hablar es lo que enseñamos a los estudiantes de física en el tercer o cuarto año del posgrado. Es mi trabajo convencerlos de que no lo eviten por no entenderlo. Mis estudiantes de física no lo entienden... eso es porque yo no lo entiendo. Nadie lo entiende”, Richard P. Feynman [Dic99].

Las ecuaciones de movimiento, acopladas a la interacción entre las partículas (via el potencial \mathcal{V}) permiten evolucionar la configuración inicial y llevarla adelante en el tiempo. En esta sección describo el formalismo teórico en que se basa la dinámica molecular.

El punto de inicio de la descripción es el principio variacional de Hamilton, que resume de manera elegante la mecánica clásica, postulando que la trayectoria en el espacio fase, que sigue un sistema mecánico es aquella para la que la integral en el tiempo $\int \mathcal{L} dt$, donde \mathcal{L} es el Lagrangiano, es un extremo.

Dado un conjunto de N coordenadas y velocidades generalizadas independientes $\{q_k, \dot{q}_k\}$ que describen un estado de un sistema conservativo (en el que todas las fuerzas son la derivada de una función \mathcal{V} llamada “de energía potencial”), de tal manera que $\mathcal{L} = \mathcal{L}(\{q_k\}, \{\dot{q}_k\}, t)$, entonces se puede probar [Gol94][Lan94][Fey87] que \mathcal{L} satisface la ecuación de Euler-Lagrange

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \left(\frac{\partial \mathcal{L}}{\partial q_k} \right) = 0. \quad (5.13)$$

Considerando la definición usual para la energía cinética \mathcal{K} y la definición de \mathcal{V} (sección 7) \mathcal{L} está definido como

$$\mathcal{L} = \mathcal{K} - \mathcal{V} = \frac{1}{2} \sum_k (m_k \dot{q}_k^2 - \mathcal{V}\{q_k\}) \quad (5.14)$$

Motivados por razones sentimentales [Fey97] quisieramos que al basar la mecánica clásica en el formalismo de Hamilton no perdiéramos la mecánica basada en la formulación de Newton, así que si consideremos un sistema con coordenadas cartesianas r_k , de la ecuación anterior obtenemos

$$m_k \ddot{\mathbf{r}}_k = \mathbf{f}_k \quad (5.15)$$

en donde m_k es la masa del átomo k y

$$\mathbf{f}_k = \nabla_{\mathbf{r}_k} \mathcal{L} = -\nabla_{\mathbf{r}_k} \mathcal{V} \quad (5.16)$$

es la fuerza sobre ese átomo. En la ecuación 5.15 reconocemos la 2ª Ley de Newton. Las otras dos leyes de Newton están contenidas en la esencia de la definición de fuerza y en la simetría que la definición de fuerza tiene para con las fuerzas que partículas i y j se ejercen mutuamente (consecuencia también de la propuesta que hacemos de la forma de \mathcal{V}). Haile [Hai92] propone que podemos obtener la 3ª Ley de la segunda proponiendo un sistema aislado de sólo dos elementos en el que, por definición, no hay fuerzas externas. La fuerza total es cero $\mathbf{F}_{total} = 0$, de esta forma la suma de las fuerzas internas debe de ser cero $\mathbf{F}_{total} = \mathbf{F}_1 + \mathbf{F}_2 = 0$, lo que implica que la fuerza que el elemento 1 ejerce sobre el 2 (\mathbf{F}_2) debe balancear a la que el elemento 2 ejerce sobre el 1 (\mathbf{F}_1), y entonces $\mathbf{F}_1 = -\mathbf{F}_2$ que es la 3ª Ley de Newton.

Habiendo recuperado la mecánica de Newton podemos dormir tranquilos sintiendo que descansamos en los hombros del gigante Newton, sin embargo, no debemos olvidar que a cambio de la seguridad hemos heredado muchos de los prejuicios y complejos acumulados por una vieja tradición.

Si estuviéramos interesados en hacer dinámica molecular de moléculas, estas ecuaciones se referirían al centro de masa de cada molécula y \mathbf{f}_k representaría la fuerza total sobre la molécula k .

Las ecuaciones para movimiento rotacional también pueden ser expresadas según la ecuación 5.13.

El momento generalizado p_k , conjugado con q_k , está definido como

$$p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \quad (5.17)$$

y las ecuaciones de movimiento en forma hamiltoniana son

$$\dot{q}_k = \frac{\partial \mathcal{H}}{\partial p_k} \quad (5.18a)$$

$$\dot{p}_k = -\frac{\partial \mathcal{H}}{\partial q_k} \quad (5.18b)$$

para el hamiltoniano definido por

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \sum_k \dot{q}_k p_k - \mathcal{L}(q, \dot{q}) \quad (5.19)$$

donde se ha asumido que podemos escribir al \dot{q}_k de la derecha como una función del momento \mathbf{p} . Para una gran cantidad de problemas, incluyendo el de frontera de grano, el potencial \mathcal{V} es independiente tanto de la velocidad como del tiempo, haciendo que \mathcal{H} sea entonces igual a la energía del sistema [Gol94, cap. 8]. En sistemas cartesianos las ecuaciones de Hamilton se convierten entonces en

$$\dot{r}_i = \frac{p_i}{m_i} \quad (5.20a)$$

$$\dot{p}_i = -\nabla_{\mathbf{r}_i} \mathcal{V} = \mathbf{f}_i \quad (5.20b)$$

Calcular las trayectorias descritas por partículas (o centros de masa de moléculas) requiere solucionar un sistema de $3N$ ecuaciones diferenciales de segundo orden (las ecuaciones 5.15) o un conjunto equivalente de $6N$ ecuaciones diferenciales de primer orden (las ecuaciones 5.20a). Antes de discutir cómo las resolvemos, es interesante hacer algunos comentarios acerca de las ecuaciones mismas.

Una consecuencia de la ecuación 5.18a o igualmente de las ecuaciones 5.13 y 5.17 es que en circunstancias especiales un momento generalizado en particular p_k puede conservarse ($\dot{p}_k = 0$). Para ello es necesario que \mathcal{L} (y \mathcal{H}) sean independientes de la coordenada generalizada conjugada q_k . Para cualquier conjunto de partículas, es posible elegir seis coordenadas generalizadas especiales, de tal manera que cambios en ellas correspondan a traslaciones y rotaciones sobre el centro de masa. Las $3N - 6$ coordenadas restantes involucran el movimiento relativo entre las partículas del sistema. En el caso a tratar en esta tesis, el potencial no depende de la orientación de las partículas, ni hay campo externo aplicado, por lo que la función de potencial \mathcal{V} depende sólo de la magnitud de la separación entre las partículas. \mathcal{V} , \mathcal{H} y \mathcal{L} son, entonces, funciones independientes de las seis coordenadas generalizadas especiales.

El momento conjugado en coordenadas cartesianas es el momento lineal total

$$\mathbf{P} = \sum_i \mathbf{p}_i \quad (5.21)$$

y el momento angular total

$$L = \sum_i \mathbf{r}_i \times \mathbf{p}_i = \sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i \quad (5.22)$$

en donde he colocado el origen de las coordenadas sobre el centro de masa del sistema.

En la práctica no se pueden tener sistemas completamente aislados, por lo que hace falta un criterio más general para estudiar los principios de conservación. Un tratamiento sumamente interesante lo hace Goldstein en el capítulo 8 de su libro de Mecánica Clásica [Gol94], basado en argumentos de simetría. Si el sistema (es decir, \mathcal{H}) es invariante respecto a traslaciones en una dirección en particular, la componente del momento correspondiente a esa dirección se conserva. Análogamente, si el sistema es invariante respecto a rotaciones sobre un eje en particular, entonces la componente del momento angular correspondiente a ese eje se conserva.

Con respecto a las leyes de conservación, ha quedado para el final la más importante: la Ley de la conservación de la energía. Es fácil mostrar que si \mathcal{V} y \mathcal{K} no dependen explícitamente del tiempo $\frac{d\mathcal{H}}{dt} = 0$, es decir, el Hamiltoniano del sistema es una constante de movimiento. Esta ley de conservación es válida sin importar si hay o no un potencial externo: la condición esencial es que no haya una dependencia explícita del tiempo (o de la velocidad) en las fuerzas que actúan en el sistema.

Respecto a las ecuaciones de movimiento es notable su reversibilidad en el tiempo, así que invertir el signo de todas las velocidades (momentos), tiene el efecto de que las partículas recorran sus trayectorias en sentido inverso. Debemos esperar que trayectorias generadas por computadora también tengan esta propiedad y será una forma de validar un programa.

La derivada espacial del potencial aparece en varias de las ecuaciones recién tratadas (por ejemplo la ecuación 5.20a), esto conlleva una diferencia cualitativa en la forma del movimiento y la manera en que se resuelven las ecuaciones, dependiendo de si \mathcal{V} es o no una función continua de la posición de las partículas. Para usar los métodos de solución por diferencias finitas que describo en la sección 6.2 es esencial que las posiciones de las partículas varíen suavemente en el tiempo. Así se podrán hacer desarrollos en serie de Taylor de $\mathbf{r}(t)$ con respecto a t . Cuando el potencial es discontinuo (cómo en el caso de los potenciales de esfera dura o de pozo rectangular de la sección 7.4.2) ocurren interacciones instantáneas entre las partículas (colisiones), en donde sus velocidades cambian en forma discontinua, cómo por efecto de un impulso instantáneo. La dinámica de las partículas en el momento de cada colisión debe tratarse en forma explícita, y por separado del movimiento suave inter-colisiones. Identificar colisiones sucesivas es una propiedad fundamental de un programa de dinámica molecular para ese tipo de sistemas (llamados *de esferas duras*) y es tratado en detalle en la sección 3.6 del Allen [All87] y en el Haile [Hai92].

Es posible restringir el movimiento de las partículas. La forma de restricción más tratada en la literatura [Rap95][Msi97] son las constricciones a los átomos para que se acomoden en moléculas y poder estudiar la dinámica de sistemas moleculares. Tanto los libros de mecánica cómo la bibliografía mencionada arriba permiten extender las ecuaciones de movimiento para tratar esta situación, y aunque es posible que las soluciones resulten un tanto complicadas, hay casos en que sus efectos se tratan sin mayor dificultad. En la sección 5.3.2 describo métodos de restricción.

5.3.1 2D ó 3D

Una de las grandes virtudes de hacer simulaciones de dinámica molecular es la posibilidad de simular sistemas con diferentes grados de libertad. No es necesario utilizar sistemas tridimensionales, podemos limitarnos a sistemas de dos o incluso una sola dimensión y la mayor parte de los cálculos sería igual. Es común hacer programas que, en forma general, puedan tratar con sistemas de diferente número de dimensiones, permitiendo estudios más completos.

Limitar las simulaciones se puede hacer de diferentes maneras:

1. Construir un sistema que exista sobre una o dos dimensiones. Si las únicas fuerzas son internas, esperaríamos que el sistema se mantuviera sin moverse en la(s) restantes dimensión(es) por el resto de la simulación. Desgraciadamente esto no necesariamente ha de ocurrir debido a los errores introducidos por la computadora. Si las dimensiones elegidas coinciden con las canónicas los errores normalmente serán menores y esperamos que el sistema se mantenga restringido por más tiempo. Este método es el preferido para sistemas simples que se puedan restringir a las dimensiones canónicas y que se simulen por tiempos cortos.
2. Modificar los algoritmos y hacer cálculos específicos para sistemas de una o dos dimensiones. Este es el método preferido para simulaciones de larga duración y/o en condiciones inestables que pudieran romper la simetría y sacar al sistema de la restricción inicial.

3. Constrañir los movimientos de los átomos a una o dos dimensiones, independientemente de la dirección de las fuerzas sobre ellos. Tiene la desventaja de que al no aplicar el efecto de las fuerzas según las ecuaciones de movimiento, estamos agregando errores de integración. Estos errores se mostrarán en fallas de conservación de la energía y otros efectos “extraños” en el comportamiento del sistema. Este método se utiliza en condiciones en las que no esperamos (o necesitamos) una conservación de la energía, y es más común en simulaciones combinadas con Monte Carlo.

No sólo se acostumbra limitar el número de dimensiones, también puede aumentarse, buscando representar en las dimensiones adicionales otras variables de interés. Por supuesto, en estos casos las ecuaciones de movimiento usuales no son aplicables.

5.3.2 Restricciones y constricciones

La diferencia esencial entre las *restricciones* y las *constricciones* es que una constricción es una restricción absoluta en los cálculos, mientras que una restricción se aplica indirectamente (por ejemplo a través de la energía) y trata de forzar los cálculos hacia cierta restricción. Restricciones y constricciones permiten enfocar el cálculo en una región o configuración de interés, así como montar experimentos. Estos experimentos son uno de los usos principales de la simulación por computadora.

Un ejemplo de restricción sería fijar los átomos en el espacio, sin permitirles moverse, mientras que en una constricción se agrega un potencial que aumenta mucho la energía de los átomos que se muevan.

Agregar restricciones y constricciones puede tener otras ventajas. por ejemplo, fijar átomos en el espacio sin permitirles movimiento alguno reduce el tiempo de cálculo de dos formas: 1) Los términos de la energía que involucran sólo átomos inmoviles no necesitan ser contabilizados, pues únicamente agregan una energía constante al total. Dado que sus posiciones no cambian sus contribuciones tampoco. 2) Fijar los átomos reduce el número de grados de libertad del sistema, disminuyendo el tiempo que toma converger al equilibrio y muestrear el espacio de configuraciones.

Es común aplicar restricciones a la distancia entre pares atómicos, agregando términos a la energía que sólo dependen de esa pareja, usualmente una función armónica

$$\mathcal{E} = K(r_{ij} - r_{objetivo})^2; \quad (5.23)$$

Otro método común es forzar a los átomos a ocupar una plantilla (*Template forcing* en inglés). Una manera de hacerlo es agregar el siguiente término a la expresión de la energía

$$\mathcal{E} = \sqrt{\sum_{\text{pares}}^N K_i \frac{(r_i - r_i^{\text{plantilla}})^2}{N}} \quad (5.24)$$

6

$$\mathcal{E} = \sum_{\text{pares}}^N K_i (r_i - r_i^{\text{plantilla}})^2 \quad (5.25)$$

La energía de la ecuación 5.24 es proporcional a la *rms* (del inglés *root mean square*) del átomo análogo en la plantilla. La ecuación 5.25 representa una restricción más simple, con una unión armónica entre átomos análogos. En cualquiera de las formas la suma es sobre una lista de las parejas de átomos a restringir, uno de la plantilla y otro móvil. La primera da la mejor *rms* por el mínimo costo energético, pero los átomos pueden mantener relativamente lejos de su posición en la plantilla. La segunda forma restringe cada átomo individualmente, dando un ajuste *rms* peor que el de la ecuación 5.24, pero donde ningún átomo se puede alejar mucho de su posición en la plantilla. La ecuación 5.25 permite también utilizar una constante K_i diferente para cada pareja.

Hay otra forma de restricción llamada *atar* (*tethering* en inglés), similar a forzar una plantilla, donde esta consiste de las posiciones iniciales.

6. EJECUTANDO, MIDIENDO Y ANALIZANDO

“Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics” David L. Goodstein.

6.1 Usando unidades más simples

Por simplicidad los programas de simulación comúnmente se escriben de tal manera que las cantidades usadas en los cálculos sean adimensionales, haciendo una transformación de unidades de valores “reales” a adimensionales o *reducidos*. Hay varias razones para hacer esto: 1) Poder trabajar con unidades cercanas a la unidad, reduciéndose la necesidad de la notación científica; 2) Usar ecuaciones simplificadas, que absorben la mayoría de los parámetros en las unidades. De esta forma, si observamos que la velocidad de un átomo es 2×10^6 inmediatamente nos damos cuenta de que debe haber un error, aún cuando sea posible que en la realidad ese átomo se moviera a $2 \times 10^6 \frac{\text{km}}{\text{hora}}$. De esta forma se puede observar la evolución de un programa y detectar los errores, ya sea en los cálculos, en el algoritmo o en la programación.

La razón más familiar para usar unidades reducidas se relaciona con la noción general de *escala*, que sugiere que un modelo puede describir toda una clase de problemas. Una vez que las propiedades de un miembro de la clase se conocen en unidades adimensionales pueden ser escaladas a las unidades físicas apropiadas para el resto. El uso de unidades reducidas también ayuda a evitar un error particularmente embarazoso: llevar a cabo simulaciones por duplicado. Desde un punto de vista estrictamente práctico, cambiar a las nuevas unidades disminuye el riesgo de encontrar valores fuera del rango representable por una computadora (o al menos reduce los errores asociados a los cálculos computacionales).

Una virtud más de las unidades reducidas es que todas aquellas unidades que al reducirse se sustituyan por la unidad, dejarán de representar cálculos en la simulación, reduciendo el tiempo de cómputo y centralizando los cambios necesarios al pasar de un sistema a otro.

No siempre se hacen las mismas transformaciones de unidades, y la elección depende de diferentes factores, como por ejemplo el potencial que se vaya a utilizar. Como ejemplo muestro una de las formas estándar, pensada para simplificar problemas que usan el potencial de Lennard-Jones (sección 7.4.3).

Lo más simple en este caso es utilizar como unidades de longitud, masa y energía a σ , m y ϵ respectivamente, haciendo las transformaciones $r^* = r/\sigma$, $m^* = 1$ y $\mathcal{V}^* = \mathcal{V}/\epsilon$ en donde $*$ designa una unidad como reducida o adimensional.

Con estas transformaciones el potencial de Lennard-Jones se transforma de la ecuación 7.28 a

$$\mathcal{V}^*(r^*) = 4 \left[\left(\frac{1}{r^*} \right)^{12} - \left(\frac{1}{r^*} \right)^6 \right] \quad (6.1)$$

y la ecuación de movimiento de Newton de 5.15 a

$$\ddot{r}_{ix} = 48 \sum_{j \neq i} \left(\frac{1}{r_{ij}^{14}} - \frac{1}{2r_{ij}^8} \right) r_{ij} \quad (6.2)$$

con expresiones análogas para y y z .

La unidad de temperatura es ϵ/k_B y dado que cada grado de libertad contribuye con $\frac{1}{2}k_B T$ a la energía cinética, la temperatura de un sistema de d -dimensiones será (sección 6.6.7)

$$T = \frac{1}{dN} \sum_i v_i^2 \quad (6.3)$$

en donde he fijado $k_B = 1$ para que la unidad de temperatura quede definida. Estrictamente, de los dN grados de libertad, d se eliminan por la conservación del momento, pero si N es grande este detalle es casi indistinguible.

Cantidades Fundamentales	
Masa	$m^* = 1$
Longitud	$r^* = r/\sigma$
Energía	$\mathcal{E}^* = \mathcal{E}/\epsilon$
Tiempo	$t^* = t\sqrt{\epsilon/m}/\sigma$
Cantidades Derivadas	
Densidad	$\rho^* = \rho\sigma^3$
Temperatura	$T^* = Tk_B/\epsilon$
Presión	$P^* = P\sigma^3/\epsilon$
Velocidad	$v^* = v\sqrt{m/\epsilon}$
Fuerza	$F^* = F\sigma/\epsilon$
Torque	$\tau^* = \tau/\epsilon$
Capacidad calorífica	$C_v^* = C_v/k_B$
Tensión superficial	$\gamma^* = \gamma\sigma^2/\epsilon$
Compresibilidad adiabática	$k_s^* = k_s\epsilon/\sigma^3$

Tabla 6.1 Unidades reducidas.

La forma reducida de otras cantidades derivadas se muestra en la tabla 6.1:

Cómo un ejemplo de unidades reducidas supongamos que simulamos argón líquido, las relaciones entre las unidades adimensionales reducidas y las físicas son las siguientes: Las longitudes se expresan en términos de $\sigma = 3.4\text{\AA}$. Las unidades de energía se especifican con $\epsilon/k_B = 120\text{ K}$, implicando que $\epsilon = 120 \times 1.3806 \times 10^{-16}\text{ erg/átomo}$. Dado que la masa de un átomo de argón es $m = 39.95 \times 1.6747 \times 10^{-24}\text{ g}$ la unidad de tiempo corresponde a $2.161 \times 10^{-12}\text{ s}$. Por ello un intervalo de $dt = 0.005$ usado en la integración de las ecuaciones de movimiento corresponde aproximadamente a 10^{-14} s . Finalmente, si N átomos ocupan una región cúbica de lado L , entonces una densidad típica de un líquido de 0.942 g/cm^3 implican que $L = 4.142N^{1/3}\text{\AA}$, que en unidades reducidas se convierte en $L = 1.218N^{1/3}$.

Transformar a unidades adimensionales evita problemas y generaliza soluciones, permitiendo poner más atención a la esencia del problema: *“La característica distintiva de un ingeniero es la rapidez con que puede reducir de pies cuadrados a acres y demás. Si su cerebro se liberara de ello, podría hacer más en otros lugares, y tener más tiempo para investigar las propiedades de la materia”*. W. Thomson (Lord Kelvin) [Tho04].

Más adelante en la sección 9.7 hago una descripción del manejo de unidades en el programa generado con esta tesis. Describo la manera en que se redujeron las unidades y la forma en que podemos convertir de unidades reales a unidades reducidas y vice versa. También abundo en el ejemplo del potencial de Lennard-Jones para argón, mostrando en detalle las transformaciones de unidades utilizadas.

6.2 Integrando las ecuaciones de movimiento

“La prueba de una inteligencia de primera clase es la capacidad de tener dos ideas opuestas en la mente al mismo tiempo y aún así poder funcionar.” F. Scott Fitzgerald.

Al encontrar soluciones para las ecuaciones de la sección 5.3 obtenemos un muestreo del espacio fase. Recordemos que solucionar estas ecuaciones no es la única opción para hacer este tipo de muestreo, sino la opción que elegimos en dinámica molecular (4.3).

La manera estándar de solucionar ecuaciones diferenciales ordinarias, cómo la 5.15, es conocida como *de diferencias finitas*, la idea general es la siguiente: Dadas las posiciones y velocidades en el instante t , intentamos extrapolarlas para un tiempo (usualmente posterior) $t+dt$, con el mayor grado de precisión posible. Las ecuaciones son solucionadas paso a paso y, por lo general, la elección del tamaño de dt depende del método de solución utilizado. La mayor parte de los métodos supone que, para intervalos de tiempo cortos, la función solución puede ser aproximada por una función simple de tratar. La precisión del método está íntimamente relacionada con cuánto se pueda mantener esta afirmación. Por lo general para los sistemas que queremos solucionar numéricamente la afirmación se mantiene sólo para dt relativamente pequeñas, donde “relativamente pequeñas” es con respecto a algún tiempo característico del sistema, por ejemplo el que le toma a una partícula realizar una oscilación alrededor de su punto de equilibrio.

Hay muchos algoritmos basados en la idea de diferencias finitas. Podemos distinguir dos formas principales de atacar el problema, la primera, y más común, es la que toman los algoritmos del tipo de Verlet, que uso en esta tesis, y que explico primero. La segunda es la de los métodos de Predictor-Corrector, por lo general un poco más complicados de implementar, pero que prometen algunas ventajas en la calidad de los resultados.

6.2.1 Euler

Probablemente la manera más simple de integrar las ecuaciones es el método de Euler. Se basa en una expansión de Taylor a primer orden:

$$\frac{dy}{dx} = f(x, y) \quad (6.4)$$

$$y_{i+1} = y_i + f(x_i, y_i)\Delta x_i \quad (6.5)$$

$$\Delta x = x_{i+1} - x_i \quad (6.6)$$

Este método tiene la desventaja de necesitar pasos muy pequeños para mantener constante la energía, ya que el método sólo es correcto hasta un orden dt , por lo tanto se reserva para la solución de problemas en los que sea posible utilizar dt pequeños, ya sea porque no es importante la rapidez de solución o porque el sistema tiene ecuaciones particularmente simples. Para una discusión más detallada del método de Euler sugiero el texto de DeVries [Dev97] y el texto de Koonin [Koo90].

6.2.2 Runge-Kutta

Runge-Kutta es una familia completa de métodos, dependiendo del orden de las aproximaciones utilizadas. Como ejemplo muestro el algoritmo más usado, el de orden 2 [Koo90]:

$$y_{n+1} = y_n + dtf(x_n + \frac{1}{2}dt, y_n + \frac{1}{2}k) + \mathcal{O}(dt^3) \quad (6.7)$$

$$k = dtf(x_n, y_n). \quad (6.8)$$

El algoritmo de Runge-Kutta ofrece cálculos mucho más precisos con un costo computacional demasiado grande, porque requiere del cálculo de las interacciones atómicas más de una vez por intervalo temporal. Es más complicado que el método de Euler y permite usar un intervalo dt mayor, sin embargo dado que la mayor parte del tiempo de simulación transcurre en el cálculo de las interacciones, no es el óptimo.

Es un método muy común en la física para resolver computacionalmente ecuaciones diferenciales, pero para simulación molecular los métodos de Verlet y de predictor-corrector descritos más adelante tienen un mejor balance de precisión vs tiempo de cálculo.

Nuevamente sugiero el texto de DeVries [Dev97] para una discusión más detallada de este método.

6.2.3 El algoritmo de Verlet

Adoptado por Verlet en 1967, es también atribuido a Störmer [Gea71] (por lo que entre los matemáticos se conoce como *algoritmo de Störmer de 3^{er} grado*). Soluciona en forma directa ecuaciones de segundo grado como la segunda ley de Newton de nuestro problema. El método parte de las posiciones $\mathbf{r}(t)$ y aceleraciones $\mathbf{a}(t)$ en el instante t y las posiciones $\mathbf{r}(t - dt)$ en el instante $t - dt$. Las posiciones de los elementos del sistema se modifican según

$$\mathbf{r}(t + dt) = 2\mathbf{r}(t) - \mathbf{r}(t - dt) + dt^2\mathbf{a}(t) \quad (6.9)$$

Esta regla de avance implica que son necesarias las posiciones en dos instantes para poder predecir una tercera en el futuro, con el efecto negativo de necesitar más espacio para almacenar los datos, o ser un poco más complicado de implementar.

De la ecuación 6.9 puede llamarnos la atención que: para predecir la posición de una partícula no es necesaria su velocidad en forma explícita, lo que aparentemente contradice la intuición, pero se explica de la siguiente forma: si hacemos los desarrollos de Taylor alrededor de $r(t)$

$$r(t + dt) = r(t) + dtv(t) + \frac{1}{2}dt^2a(t) + \dots \quad (6.10)$$

$$r(t - dt) = r(t) - dtv(t) + \frac{1}{2}dt^2a(t) - \dots \quad (6.11)$$

sumamos ambas ecuaciones y despejamos, obtenemos la ecuación 6.9. Que las velocidades no sean necesarias para calcular las posiciones de las partículas no las hace inútiles, mantienen su carácter de indispensables porque en ellas hay información de la energía cinética de la partícula y de la temperatura del sistema completo. Las velocidades se obtienen fácilmente recordando que (como dt es pequeño)

$$v(t) = \frac{r(t + dt) - r(t - dt)}{2dt} \quad (6.12)$$

Mientras la ecuación 6.9 implica errores de magnitud máxima dt^4 en la predicción local de la posición, las velocidades obtenidas con la ecuación 6.12 implican errores de orden dt^2 . Es posible estimar $v(t)$ con mayor precisión, pero esto sólo complica el método y tiene ventajas poco notables. El método tiene una complicación natural: el valor de la velocidad en el instante t sólo lo conocemos una vez que conocemos la posición en el instante $t + 1$; esta "diferencia de fase" ente ambos puede complicar la programación del algoritmo, su uso e interpretación. Por lo general se prefiere utilizar variaciones de este algoritmo que eliminen este problema, de las que hablare un poco más adelante.

Una virtud de las ecuaciones 6.9 y 6.12 del algoritmo de Verlet es que son centradas y simétricas respecto a t , por lo que permiten invertir el tiempo. Otra virtud del método de Verlet se vera con facilidad más adelante al comparar con los métodos de tipo Predictor-Corrector (PC), pues el método de Verlet permite avanzar las posiciones en un sólo paso, y los basados en PC requieren de al menos dos.

La implementación del algoritmo de Verlet requiere almacenar en memoria $9N$ valores numéricos (que más tarde veremos es muy poco comparado con otros métodos). Además es simple de programar, exactamente reversible en el tiempo y, dadas fuerzas conservativas, permite conservar el momento lineal bien. El método presenta una buena conservación de la energía, incluso para intervalos de tiempo largos (ver la sección 6.2.12)). Por otro lado, la manera de tratar las velocidades es un tanto rara, y la forma del algoritmo introduce errores. por ejemplo, en la ecuación 6.9 un término ($\mathcal{O}(dt^2)$) es agregado a la resta de dos términos grandes ($\mathcal{O}(dt^0)$), para generar la trayectoria.

Una imprecisión del algoritmo surge de la manera en que calculamos las velocidades en la ecuación 6.12, en donde restamos dos posiciones separadas por un intervalo muy pequeño ($2dt$), que seguramente son parecidas y su diferencia muy pequeña. Una vez obtenida la pequeña diferencia la dividimos entre otro número pequeño, lo que tiene el efecto de amplificarla. Como ejemplo ilustrativo supongamos que estamos utilizando una computadora que nos permite 6 dígitos de precisión. Si tenemos que una partícula ha avanzado muy poco durante el tiempo $2dt$, sus posiciones pueden ser 0.100000 y 0.100100 la diferencia entre ellas es 0.000100 que tiene 6 dígitos, pero cuyos primeros dígitos siempre serán cero y dependiendo de lo poco que haya avanzado la partícula (ya sea por una velocidad pequeña o un dt muy pequeño) habrá más o menos dígitos cero; Después dividimos ese número entre $2dt$, también pequeño, al menos 0.02, y encontramos que la velocidad es 0.00500 ¿cual es el problema?, que dividir entre 0.02 tiene el efecto de multiplicar por 50, de tal manera que los dos últimos dígitos del resultado no pueden ser otros que cero. El resultado tendrá muchos ceros, todos los que son efecto de lo similares que son las posiciones en dos instantes tan cercanos en el tiempo más los que son producto de multiplicar por un número grande entero. Curiosamente, y a diferencia de otras imprecisiones, los problemas aumentan cuando disminuye el valor de dt . Aparece así un efecto interesante: *la precisión de un algoritmo no depende unicamente de la consistencia numérica de las ecuaciones, el error que estas introducen y la precisión numérica de la máquina, sino también de otros factores difíciles de calcular.* El resultado neto es muy importante: Conforme disminuye dt el error asociado a un algoritmo disminuye, hasta llegar a un mínimo, para luego aumentar nuevamente. Dando lugar a un dt óptimo. El óptimo es una función del algoritmo, pero también de la precisión de la máquina y de otros factores. En ocasiones es necesario hacer repetidas pruebas hasta identificar el óptimo del problema en cuestión.

Es posible evitar algunos problemas cambiando de algoritmo, y por ello existen variantes del método de Verlet, siendo los más populares la de "Salto de Rana" y la de "Verlet Velocidad" que describo a continuación, y que han provocado que el algoritmo de las ecuaciones 6.9 y 6.12 se conozca como de Verlet Simple.

6.2.4 El algoritmo Salto de Rana

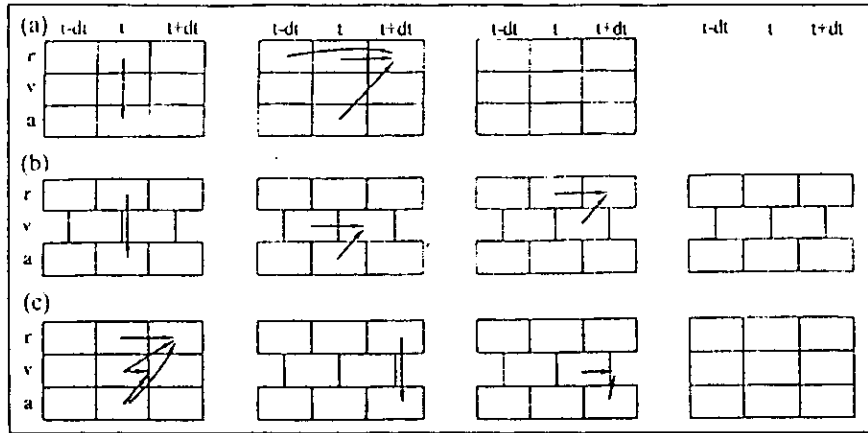


Figura 6-1 Diferentes formas del algoritmo de Verlet. (a) Método original de Verlet. (b) Método de Salto de rana. (c) Verlet Velocidad. Se muestran pasos sucesivos en la implementación de cada algoritmo. [All87]

Salto de Rana es un nombre curioso para un algoritmo. Su origen quedara claro al conocer el algoritmo

$$r(t + dt) = r(t) + dtv(t + \frac{1}{2}dt) \quad (6.13)$$

$$v(t + \frac{1}{2}dt) = v(t - \frac{1}{2}dt) + dta(t) \quad (6.14)$$

La forma de calcular las posiciones se deriva directamente de la mecánica, suponiendo que la velocidad en el instante intermedio $t + \frac{1}{2}dt$ es constante en todo el intervalo $[t, t + dt]$. Sólo es necesario almacenar en memoria las posiciones $r(t)$, las aceleraciones $a(t)$ y las velocidades $v(t - \frac{1}{2}dt)$ en un instante $t - \frac{1}{2}dt$ intermedio. La ecuación 6.13 se calcula primero, luego las velocidades "saltan" las coordenadas y calculamos las velocidades en el instante $t + \frac{1}{2}dt$, obteniendo $v(t + \frac{1}{2}dt)$. Ya con las velocidades, en instantes $\frac{1}{2}dt$ antes y después que las posiciones, las interpolamos linealmente para encontrar el valor de la velocidad al instante t :

$$v(t) = \frac{1}{2} \left[v(t + \frac{1}{2}dt) + v(t - \frac{1}{2}dt) \right] \quad (6.15)$$

A diferencia del algoritmo Simple de Verlet, las velocidades aparecen en forma explícita en el cálculo de las posiciones. La razón de calcular la velocidad al instante t es para completar los valores y poder calcular propiedades con posiciones y velocidades en el mismo instante.

El orden de los cálculos es el siguiente

$$\text{velocidades}(t + \frac{1}{2}dt) \rightarrow \text{posiciones}(t + dt) \rightarrow \text{aceleraciones}(t + dt) \rightarrow \text{velocidades}(t + \frac{3}{2}dt) \quad (6.16)$$

El diagrama de la figura 6-1 compara los métodos.

Ya no es necesario calcular la resta de dos cantidades muy parecidas, ahora se suman, mejorando la precisión en los cálculos. Es necesario más espacio en memoria para almacenar los datos (sección A.1.2).

El algoritmo de Salto de Rana es simple y mejora la precisión con respecto al Simple de Verlet, pero aún deja mucho que desear en cuanto al manejo de las velocidades.

6.2.5 Algoritmo de Verlet Velocidad

Swope, Andersen, Berens y Wilson propusieron en 1982 [Swo82] un algoritmo equivalente al de salto de rana, que minimizando los errores de redondeo, almacena las posiciones, velocidades y aceleraciones para el mismo instante t : este algoritmo conocido como de Verlet Velocidad es

$$r(t + dt) = r(t) + dtv(t) + \frac{1}{2}dt^2a(t) \quad (6.17)$$

$$v(t + dt) = v(t) + \frac{1}{2}dt \{a(t) - a(t + dt)\} \quad (6.18)$$

Este algoritmo es preferible por la simplicidad, claridad y eficiencia.

El algoritmo sólo requiere almacenar las posiciones, velocidades y aceleraciones, pero hay dos diferentes maneras de implementarlo dependiendo de la memoria de que se disponga. Si tenemos memoria disponible almacenamos además un segundo valor para las aceleraciones para poder calcular el promedio de la aceleración en el intervalo $[t, t + dt]$. Si queremos ahorrar memoria podemos calcular la velocidad intermedia

$$v(t + \frac{1}{2}dt) = v(t) + \frac{1}{2}dta(t) \quad (6.19)$$

cuando sólo hemos calculado las aceleraciones en el instante t . Después calculamos las aceleraciones al instante $t + dt$, sustituyendo las anteriores, y terminamos el avance en la velocidad con

$$v(t + dt) = v(t + \frac{1}{2}dt) + \frac{1}{2}dta(t + dt) \quad (6.20)$$

Ahora podemos calcular la energía cinética del instante $t + dt$.

En 6.17 podemos ver que la ecuación para avanzar las posiciones en el tiempo corresponde con la que nos enseñaron en la preparatoria para el movimiento uniformemente acelerado.

Podemos comparar los diferentes algoritmos de Verlet más fácilmente con la figura 6-1. Nuevamente se necesitan $9N$ variables reales almacenadas en memoria, "pero son la estabilidad, conveniencia y simplicidad lo que lo hacen, quizás, el método más atractivo propuesto hasta hoy" dice Allen en la página 82 de su libro [All87].

Más adelante en la sección 6.2.7 veremos los llamados métodos de Predictor-Corrector. Un Predictor-Corrector de tres valores comprende el caso del Verlet Velocidad cuando lo que llamaremos el coeficiente de corrección de posición sea cero.

6.2.6 Un poco más de Verlet

Antes de dar por terminado los algoritmos de Verlet vale la pena agregar una variante más propuesta por Beeman [Bee76], Sangster y Dixon [San76] e Eastwood y Hockney [Eas80]. Consiste en lo siguiente:

$$r(t + dt) = r(t) + dtv(t) + \frac{2}{3}dt^2a(t) - \frac{1}{6}dt^2a(t - dt) \quad (6.21)$$

$$v(t + dt) = v(t) - \frac{1}{6}dta(t - dt) + \frac{5}{6}dta(t) + \frac{1}{3}dta(t + dt) \quad (6.22)$$

El método requiere de memoria suficiente para almacenar $r(t)$, $v(t)$, $a(t)$ y $a(t - dt)$. En contraparte con la complejidad de las formulas y la necesidad de almacenar las aceleraciones anteriores, es un algoritmo más preciso para las velocidades que Verlet Velocidad (6.12) y consecuentemente permite observar una mejor conservación de la energía.

No hay que olvidar que todos los métodos de Verlet presentados son en esencia equivalentes y son derivables unos de otros.

6.2.7 Los métodos Predictor-Corrector

Los métodos del tipo Predictor-Corrector propuestos en [Gea71],[Bee76],[Ber86] son métodos multivalores, en el sentido de que hacen uso de diferentes datos calculados uno o más intervalos de tiempo antes. En las dos formas más familiares [Rap95, pág. 59] podemos elegir entre usar las aceleraciones calculadas en una serie de tiempos anteriores según lo propuesto por Adams, conocido como método multipaso, o usar algunas derivadas de la aceleración en el mismo instante según el método de Nordsieck. Para métodos precisos hasta un orden δt , se puede probar que ambas formas son algebraicamente equivalentes. Son métodos de mayor orden que el de salto de rana, aunque requieren más espacio en memoria.

La manera de aplicar el método para resolver problemas de dinámica molecular es: primero usar las fórmulas de predicción para las posiciones y velocidades, luego utilizar estas predicciones para calcular las fuerzas (y aceleraciones), para finalmente corregir las posiciones y las velocidades mediante las fórmulas de corrección.

Coeficientes correctores de Gear para ecuaciones de primer orden						
Orden	C ₀	C ₁	C ₂	C ₃	C ₄	C ₅
3	5/12	1	1/2			
4	3/8	1	3/4	1/6		
5	251/720	1	11/12	1/3	1/24	
6	95/288	1	25/24	35/72	5/48	1/120

Coeficientes correctores de Gear para ecuaciones de segundo orden						
Orden	C ₀	C ₁	C ₂	C ₃	C ₄	C ₅
3	0	1	1			
4	1/6	5/6	1	1/3		
5	19/120	3/4	1	1/2	1/12	
6	3/20	251/360	1	11/18	1/6	1/60

Tabla 6.2 Coeficientes correctores de Gear (All87)

6.2.8 Métodos que utilizan variables de varios instantes

Comencemos por los métodos de multipaso (valores de varios instantes) porque las derivadas de la aceleración no son naturalmente utilizadas en la mecánica newtoniana y por claridad es mejor no enredarnos con ellas desde el principio.

Recordemos que el problema consiste en resolver una ecuación de segundo orden del tipo

$$\ddot{r} = f(r, \dot{r}, t) \quad (6.23)$$

Denotemos con superíndices p y c las formulas utilizadas en las etapas de predicción y de corrección respectivamente. El paso que predice la posición al tiempo $t + dt$ mediante una extrapolación de los valores en tiempos anteriores es

$$r^p(t + dt) = r(t) + dt\dot{r}(t) + dt^2 \sum_{i=1}^{k-1} \alpha_i f(t + [1 - i] dt) \quad (6.24)$$

Este método es conocido como de Adams-Bashforrth y para un valor de k dado genera resultados exactos para $r(t) = t^q$ siempre que $q \leq k$; en el caso general los errores locales son $\mathcal{O}(dt^{k+1})$. Notese que la ecuación 6.24 contiene la misma información que una expansión en serie de Taylor. Para que el error antes dicho sea el correcto los coeficientes $\{\alpha_i\}$ deben satisfacer el conjunto de $k - 1$ ecuaciones

$$\sum_{i=1}^{k-1} (1 - i)^q \alpha_i = \frac{1}{(q + 1)(q + 2)}, \quad q = 0, \dots, k - 2 \quad (6.25)$$

Este y otros conjuntos de ecuaciones lineales han sido resueltos y la tabla 6.2 [All87] muestra los valores de los coeficientes (todos ellos fracciones racionales).

Para las velocidades existe una ecuación equivalente

$$dt\dot{r}^p(t + dt) = r(t + dt) - r(t) + dt^2 \sum_{i=1}^{k-1} \alpha'_i f(t + [1 - i] dt) \quad (6.26)$$

con coeficientes que satisfacen las ecuaciones

$$\sum_{i=1}^{k-1} (1 - i)^q \alpha'_i = \frac{1}{(q + 2)}, \quad q = 0, \dots, k - 2 \quad (6.27)$$

Después de calcular el valor de $f(t + dt)$ usando los valores predichos para x y \dot{x} , las correcciones se hacen con la ayuda del método de Adams-Moulton [Rap95], originalmente formulado como un método implícito separado, después adoptado para refinar las estimaciones. Funciona así:

$$r^c(t + dt) = r(t) + dt\dot{r}(t) + dt^2 \sum_{i=1}^{k-1} \beta_i f(t + [2 - i] dt) \quad (6.28)$$

$$dt\dot{r}^c(t + dt) = r(t + dt) - r(t) + dt^2 \sum_{i=1}^{k-1} \beta'_i f(t + [2 - i] dt) \quad (6.29)$$

y los coeficientes cumplen con

$$\sum_{i=1}^{k-1} (2-i)^q \beta_i = \frac{1}{(q+1)(q+2)}, \quad q = 0, \dots, k-2 \quad (6.30)$$

$$\sum_{i=1}^{k-1} (2-i)^q \beta_i = \frac{1}{(q+2)}, \quad q = 0, \dots, k-2 \quad (6.31)$$

Podemos notar que los valores predichos no aparecen en la fórmula de corrección, excepto en donde se usan para evaluar f . Los coeficientes $(\alpha_1, \alpha_2, \dots, \alpha_k)$ obtenidos al resolver las ecuaciones para $k = 4$ se incluyen en las funciones de integración, y están tabulados en la tabla 6.2.

El orden en que se calculan las variables determina la cantidad de cálculos a hacer, por ejemplo, en el algoritmo que acabo de presentar las aceleraciones se calculan usando los valores predichos de las posiciones y sin corregir, así que las propiedades calculadas de las interacciones de los átomos, tendrán alguna diferencia con los que pudieran obtenerse de los valores corregidos. Hay diferentes opciones para reducir este error, la primera es volver a calcular las interacciones después de aplicar las correcciones, pero el costo computacional (en tiempo sobre todo) es generalmente inaceptable, pues la mayor parte del tiempo de simulación transcurre precisamente en el cálculo de las interacciones. También se ha probado aplicar las correcciones más de una vez, lo cual es un poco más eficiente pero aún no satisfactorio. Según Rappaport [Rap95, pág. 62] ninguna de las soluciones realmente resulta en una mejora notable en la precisión y por lo tanto no se usan.

Resolver $3N$ ecuaciones diferenciales de segundo orden no es la única manera de encontrar una solución a la ecuación 6.23. Podemos preferir resolver $6N$ ecuaciones diferenciales de primer orden (ver sección 5.3). Para estos casos también hay algoritmos de tipo Predictor Corrector, como por ejemplo para resolver la ecuación

$$\dot{r} = f(r, t) \quad (6.32)$$

las fórmulas de predicción y corrección de la posición son

$$r^p(t + dt) = r(t) + dt \sum_{i=1}^k \alpha_i f(t + [1 - i] dt) \quad (6.33)$$

$$r^c(t + dt) = r(t) + dt \sum_{i=1}^k \beta_i f(t + [2 - i] dt) \quad (6.34)$$

con coeficientes que satisfacen

$$\sum_{i=1}^k (1-i)^q \alpha_i = \frac{1}{(q+1)}, \quad q = 0, \dots, k-2 \quad (6.35)$$

$$\sum_{i=1}^k (2-i)^q \beta_i = \frac{1}{(q+1)}, \quad q = 0, \dots, k-2 \quad (6.36)$$

Estos coeficientes también están tabulados en la tabla 6.2.

6.2.9 Métodos que utilizan derivadas de alto orden

De los métodos que utilizan derivadas de la posición de mayor orden que la aceleración, el ejemplo más común es el algoritmo de Gear [All87, pág. 74][Hai92]. En él se hace una expansión de Taylor sobre el tiempo de las variables que nos interesa ver evolucionar

$$\mathbf{r}^p(t + dt) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \frac{1}{6} \delta t^3 \mathbf{b}(t) + \dots \quad (6.37)$$

$$\mathbf{v}^p(t + dt) = \mathbf{v}(t) + \delta t \mathbf{a}(t) + \frac{1}{2} \delta t^2 \mathbf{b}(t) + \dots \quad (6.38)$$

$$\mathbf{a}^p(t + dt) = \mathbf{a}(t) + \delta t \mathbf{b}(t) + \dots \quad (6.39)$$

$$\mathbf{b}^p(t + dt) = \mathbf{b}(t) + \dots \quad (6.40)$$

Estas ecuaciones no pueden generar trayectorias correctas, pues no hemos introducido aún las ecuaciones de movimiento. Esto se hace en las ecuaciones de corrección. Usualmente las ecuaciones se truncan en los términos que aparecen explícitamente. Al igual que en los métodos multipaso usamos las posiciones predichas para calcular las interacciones entre los átomos y de allí las aceleraciones corregidas $\mathbf{a}^c(t + \delta t)$, que podemos comparar con las aceleraciones y estimar el error en la predicción

$$\Delta \mathbf{a}(t + \delta t) = \mathbf{a}^c(t + \delta t) - \mathbf{a}^p(t + \delta t) \quad (6.41)$$

Este error, y los resultados del paso de predicción son las entradas para el paso de corrección

$$\mathbf{r}^c(t + dt) = \mathbf{r}^p(t + \delta t) + \alpha_0 \Delta \mathbf{a}(t + \delta t) \quad (6.42)$$

$$\mathbf{v}^c(t + dt) = \mathbf{v}^p(t + \delta t) + \alpha_1 \Delta \mathbf{a}(t + \delta t) \quad (6.43)$$

$$\mathbf{a}^c(t + dt) = \mathbf{a}^p(t + \delta t) + \alpha_2 \Delta \mathbf{a}(t + \delta t) \quad (6.44)$$

$$\mathbf{b}^c(t + dt) = \mathbf{b}^p(t + \delta t) + \alpha_3 \Delta \mathbf{a}(t + \delta t) \quad (6.45)$$

La idea es que las variables calculadas con las ecuaciones anteriores son ahora mejores aproximaciones a las variables "reales". Gear en 1966 y 1971 [All87, pág. 75] discute la manera de encontrar valores de los coeficientes α_i que permitan mayor estabilidad y precisión en las trayectorias. Si se usan más o menos derivadas de la posición en las ecuaciones se necesitarán entonces diferentes valores para los coeficientes; en la tabla 6.2 doy algunos ejemplos.

Es posible aplicar más de una vez las fórmulas de corrección, como una iteración, para refinar los resultados. Según Allen hay ocasiones en que esta iteración es la clave para obtener soluciones precisas. En teoría la predicción no tiene porque ser muy buena, esperando obtener el valor útil al aplicar las correcciones, pero en dinámica molecular, como ya hemos comentado, la mayor parte del tiempo de cálculo se usa en calcular interacciones, y cómo esto se requiere en cada paso de corrección, sería demasiado caro en computo aplicar muchas iteraciones de corrección. así que usualmente sólo se utiliza una (ocasionalmente dos), lo que eleva las expectativas de precisión en las predicciones.

6.2.10 Comparación de los algoritmos de Verlet y de Predictor-Corrector

Por la mayor flexibilidad y capacidad de obtener mejor precisión local (al menos en potencia), los métodos de Predictor-Corrector se adaptan mejor a problemas complejos, cómo los de cuerpos rígidos o dinámicas con restricciones, en donde la precisión que se pueda lograr en cada paso puede ser determinante.

Los métodos de Verlet por lo general son más fáciles de programar y necesitan menos memoria, pero tienen la desventaja de que normalmente necesitan adaptarse al problema específico a resolver. Verlet Velocidad presenta una excelente conservación de la energía, inclusive para potenciales fuertemente divergentes cómo los de Lennard Jones, para dt relativamente grandes, y debido a que puede ser programado para requerir un mínimo de memoria, es excelente para simulaciones a muy gran escala (ver sección 6.2.12).

6.2.11 Otros métodos

Hay muchos otros métodos propuestos, aunque por lo general más complicados y específicos, pero vale la pena comentar al menos la existencia de un método propuesto por Toxvaerd [Tox82], y examinado en detalle por Heyes y Singer [Hey82]. El objetivo principal de este método es calcular con la máxima precisión posible las trayectorias de las partículas, tomando en cuenta para el cálculo del movimiento de una partícula el de las partículas cercanas. Desgraciadamente el método requiere dos cálculos de la fuerza sobre las partículas para cada intervalo de tiempo. Los logros son interesantes, consiguiéndose una conservación de la energía dos o tres ordenes de magnitud mejores que los métodos de Verlet [All87, pág. 78], pero en un programa que corre tres veces más lento.

6.2.12 Eligiendo un método

Describí ya muchos métodos diferentes, explicado ventajas y desventajas de cada uno, pero aún quedan las dudas ¿cómo elegir el método más adecuado?. ¿que virtudes y defectos resultan ser los mas importantes?.

Podemos comenzar observando cuales son algunas de las cualidades que esperamos tenga un algoritmo para considerarlo exitoso (ver sección 8).

1. Ser rápido, requiriendo pocos cálculos y que estos sean los más simples para una computadora (sección 8).
2. Requerir poca memoria, es decir, almacenar solamente las variables que nos interesan al instante en que nos interesan (sección A.1.2) .
3. Permitir el uso de un intervalo δt grande, para poder simular eventos de mayor duración (sección 6.2.13).
4. Duplicar las trayectorias predichas por la mecánica clásica lo mejor posible (sección 4.6.5).
5. Satisfacer las ecuaciones de conservación para el momento y la energía, además de tener reversibilidad en el tiempo (secciones 5.3 y 6.6).
6. Simple de programar y modificar (sección 9).
7. Permitir simular un sistema lo más parecido posible a un sistema real (sección 4.6).

Cada quien decide, para el caso específico que trate, el peso que da a cada punto, y el algoritmo que más le satisface.

La velocidad es un elemento clave en un programa, y por lo general determina la cantidad y calidad de los resultados obtenidos. En la sección 8 trato los métodos para acelerar los cálculos (o acortar el tiempo de espera).

En los comienzos de la dinámica molecular la memoria era una consideración importante. Hoy en día con computadoras mucho más poderosas la memoria sólo es una consideración importante en simulaciones extremas (muchos elementos, mucho tiempo de simulación, muchos parámetros. etc).

Podemos usar un intervalo dt grande es útil para hacer más rápida la simulación y permitir simular sistemas grandes o largos períodos de tiempo, pero a la vez esta relacionado directamente con la estabilidad y precisión en las trayectorias. Usualmente se elige la precisión que queremos en las trayectorias y en las leyes de conservación, eligiendo así el dt más grande posible.

¿Que tan importante resulta reproducir las trayectorias clásicas y cumplir con las leyes de conservación?. La precisión y estabilidad de los algoritmos se mide con los errores locales y globales (sección D.2.1). Considerando las teorías de sistemas dinámicos y el caos (sección 6.2), cualquier perturbación en las variables del sistema, incluso las inevitables provocadas por el redondeo de la máquina ha de modificar la trayectoria seguida y tarde o temprano la llevará lejos de la original. La figura 6-2 tomada del libro de Allen [All87] muestra cómo una pequeña perturbación aplicada al inicio lleva tarde o temprano (unos cuantos cientos de pasos de integración) a las trayectorias a diverger. La figura muestra un promedio de la diferencia entre las posiciones de los átomos al tiempo t y las posiciones iniciales ($t = 0$). En la parte derecha de la figura 6-2 se muestra el porcentaje de la diferencia entre las energías cinéticas del sistema con la perturbación y el sistema de referencia. Podemos ver cómo las energías cinéticas se mantienen relativamente iguales durante un tiempo que depende de la perturbación inicial. Sólo nos interesa que las trayectorias sean similares durante un tiempo característico de los parámetros que queremos medir.

Si no podemos tener trayectorias exactas, tratar de que la simulación cumpla con las leyes de conservación resulta más importante, porque nos da una posibilidad de observar cuando las trayectorias se alejan de las clásicas.

No queda claro qué es mejor para mejorar la precisión, minimizando el costo en tiempo, si hacer más iteraciones en los ciclos correctores, reducir el tamaño del salto temporal dt , usar otro algoritmo o algo más, pero aparentemente cualquiera de ellas logra generar trayectorias igualmente útiles.

6.2.13 Eligiendo el tamaño de dt

Hemos visto lo clave que resulta la elección del parámetro de integración dt . Para aprovechar al máximo el tiempo de cómputo, se prefiere un intervalo lo más largo posible. sin embargo, un intervalo demasiado grande causa inestabilidad e imprecisión del algoritmo de integración (sección 5.1.2).

El intervalo a usar depende del sistema y del método de integración, como vimos en la sección 6.2.12. La principal limitación impuesta por el sistema es la máxima frecuencia de vibración a considerar. Un período de vibración debe partirse en, al menos, 8 a 10 segmentos para que un sistema atómico satisfaga las suposiciones de Verlet de que las aceleraciones son constantes en el intervalo. Por ejemplo, un sistemas moleculares en el que la frecuencia de vibración más alta sea de unos 10^{-14} s. el paso de integración debe ser de 10^{-15} s. es decir 1 fs.

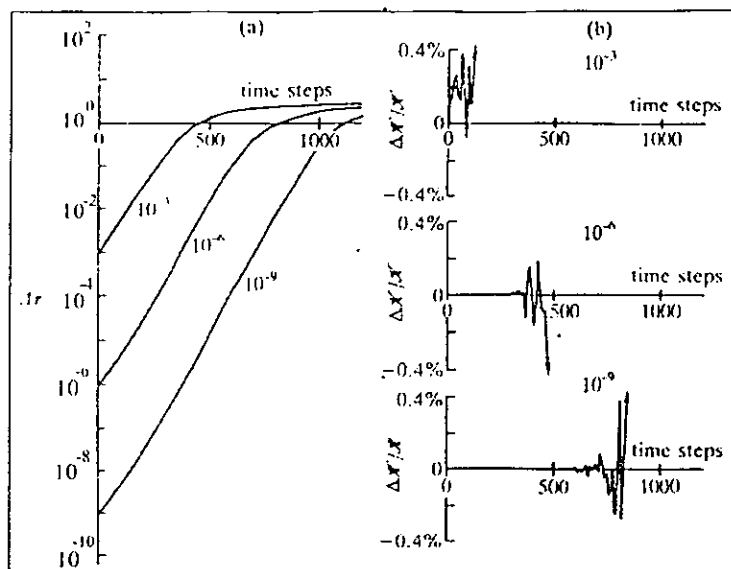


Figura 6-2 La divergencia de trayectorias en dinámica molecular. Los átomos interactúan según un potencial de Lennard-Jones. Las curvas están etiquetadas según el desplazamiento inicial en unidades de σ . a) Δr es la separación en el espacio fase entre las trayectorias perturbadas y de referencia. b) $\Delta K/K$ es el porcentaje de diferencia de las energías cinéticas. [All87]

Cuando no tenemos frecuencias vibracionales tan grandes podemos usar intervalos mayores, con sus respectivas ventajas.

Si 1 fs es adecuado para un algoritmo de Verlet, un algoritmo de Runge-Kutta requiere unos 0.05 fs y uno de Adams-Bashforth de orden 4 requiere 0.5 fs. ambos menores. con la correspondiente pérdida en eficiencia.

La figura 6-3 muestra la deriva en la energía para diferentes valores de dt y dos métodos de integración: Predictor Corrector y Salto de Rana.

La figura 6-4 muestra las fluctuaciones en la energía para diferentes algoritmos. Los cálculos son del mismo sistema t bajo las mismas condiciones iniciales. El tiempo simulado es el mismo, y las diferencias en dt han sido compensadas por el número de pasos simulados. Es interesante resaltar de la figura 6-4 que las líneas se intersectan, lo que significa que no hay un algoritmo óptimo para cualquier valor de dt . Además de la dependencia en dt al elegir el algoritmo no debemos olvidar que las fluctuaciones dependen también del potencial y del sistema.

6.3 Condiciones de frontera

Dependiendo del sistema simulado existen diferentes maneras de tratar las condiciones de frontera.

6.3.1 Libres o de vacío

Son las más simples, porque no requieren de ningún trabajo especial. Son las adecuadas si el sistema se supone finito y aislado. Hay un problema asociado a ellas: Dado un sistema, es posible calcular la energía de escape para alguno de los átomos. A pesar de que inicialmente ninguno de ellos tenga una energía mayor a esta energía de escape, durante la simulación las fluctuaciones estadísticas en la energía de cada elemento pueden provocar que alguno (o varios) átomos tengan energía suficiente y escapen a infinito. Esto tiene diferentes consecuencias: Los átomos que escapen se llevan energía, que puede dificultar mucho el cálculo de propiedades del cluster. Alejándose poco a poco estos átomos complican el estudio del sistema, pero sobre todo la visualización. Pueden hacer imposible llevar a cabo transiciones buscadas. Modifican el sistema investigado, reduciendo el número de partículas interactuando.

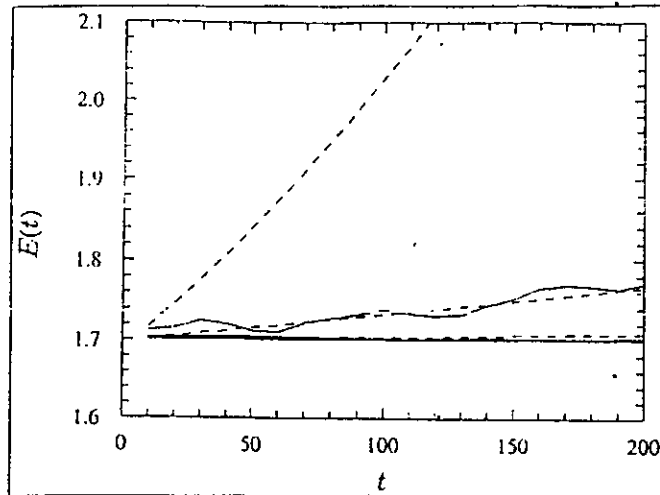


Figura 6-3 Deriva en la energía para diferentes valores de dt , para los algoritmos de Salto de Rana (línea continua) y Predictor Corrector (línea discontinua). Las primeras cuatro del Salto de Rana son para $dt = 0.00125, 0.0025, 0.005, 0.01$ y 0.02 , pero las cuatro primeras son indistinguibles. Para el PC se muestran $dt = 0.00125, 0.0025$ y 0.005 . [Rap95]

La probabilidad que algún átomo obtenga suficiente energía como para escapar del cluster disminuye conforme aumenta el tamaño del cluster y aumenta al aumentar la energía total. La manera más común de solucionar el problema es utilizar condiciones de frontera limitadas, o configuraciones iniciales de muy baja energía.

6.3.2 Para simular sistemas infinitos

Por las limitaciones de memoria y tiempo, hay un límite en el número de átomos con que se puede trabajar en una computadora. Diferentes métodos tienen diferentes limitaciones, pero de cualquier manera el número de átomos es muy limitado comparado con los átomos en muchos sistemas reales, por lo que es necesario agregar ingredientes a la fórmula para poder calcular propiedades que dependen del número de átomos.

Podemos preguntarnos si existe algún tamaño crítico a partir del cual las propiedades sean, al menos en esencia, iguales a las del sistema infinito, o si se puede observar alguna tendencia de estas propiedades al ir aumentando el tamaño del sistema, de tal manera que en el límite infinito se observen las propiedades de sistemas reales. El problema ha sido estudiado por los físicos desde hace ya mucho tiempo, y existen soluciones interesantes; una muy popular es suponer que si estudiamos un cristal desde el punto de vista atómico, el cristal es infinito, está formado por un arreglo regular (o cuasiregular) que cubre todo el espacio.

Al suponer un sistema infinito ocurre un cambio importante en el sistema: desaparecen las fronteras. Las propiedades pueden ser diferentes al menos en esas regiones.

El efecto de las fronteras del sistema puede ser muy corto (unos cuantos Å), o de muy largo alcance (todo el sistema) aunque depende de la geometría específica. En sistemas cristalinos como el que nos interesa en esta tesis, cuando el sistema aumenta de tamaño la proporción de átomos en las fronteras disminuye. Para entender porque ocurre esto podemos observar que la superficie de una esfera aumenta proporcionalmente al cuadrado del radio ($S \propto r^2$) mientras que el volumen aumenta proporcional al cubo del radio ($V \propto r^3$), por lo que si la densidad es constante y el sistema se asemeja geoméricamente a una esfera, entonces la proporción entre superficie y volumen disminuye como el radio

$$\frac{S}{V} \propto \frac{1}{r}. \tag{6.46}$$

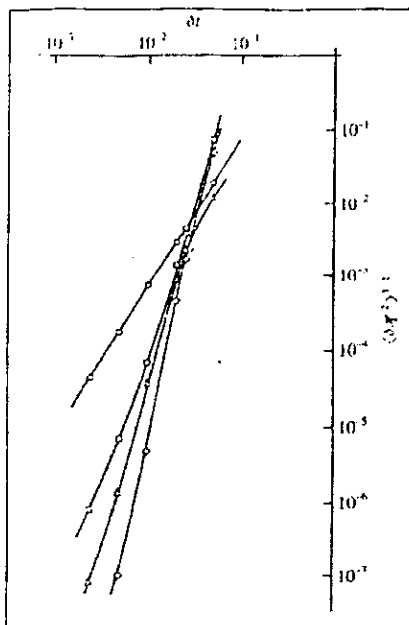


Figura 6-4 Conservación de la energía de varios algoritmos. El sistema estudiado es el mismo de la figura. En la gráfica las fluctuaciones RMS de la energía $\langle \delta H^2 \rangle^{1/2}$ para varias corridas comenzando con las mismas condiciones iniciales y durante el mismo tiempo de simulación t , pero usando diferentes intervalos dt y número de pasos. La gráfica es en escala log-log. Las curvas corresponden a los algoritmos de Verlet (círculos), Gear de 4° orden (cuadros), Gear de 5° orden (triángulos) y Gear de 6° orden (diamantes). Notese que las rectas se cruzan, lo que muestra que no hay un algoritmo óptimo para cualquier tamaño de dt . [All87]

6.3.3 Condiciones de frontera periódicas

Existen diferentes maneras de sobreponerse a este problema. la más común es implementar condiciones de frontera periódicas (PBC por las siglas en inglés de *Periodic Boundary Conditions*) propuestas por Born y von Karman en 1912 [Bor12].

Para usar condiciones de frontera periódicas en un sistema con N átomos confinados a un volumen V , imaginamos que el volumen V es sólo una pequeña parte del volumen total del material. Al volumen V le llamamos *celda original* (o *celda primaria*), y supondremos que es representativo del material en general, de tal manera que el material completo se supone compuesto por la celda original y un número infinito de replicas de ella a las que llamaremos *celdas imagen*. Las celdas imagen son, por construcción, idénticas en tamaño, forma y todas las propiedades a la celda original. La idea es que la celda original se repita periódicamente en todas las direcciones, cubriendo todo el espacio.

Suponer que todas las celdas son idénticas y equivalentes tiene consecuencias particularmente interesantes. por ejemplo, si por alguna razón una partícula se aleja mucho de las demás, de tal manera que abandona el volumen de la celda original, entonces en cada celda imagen habra otra partícula abandonando, y dado que las celdas teselan el espacio, mientras por un lado de la celda original una partícula sale, por el otro entra otra partícula idéntica proveniente de la celda imagen de al lado. En la figura 6-5 [All87, pág. 24] se observa una versión bidimensional de un sistema periódico como el propuesto, con la celda original en el centro y las celdas imagen marcadas A,B,C, etc. en forma arbitraria. Cuando la partícula 1 abandona la celda original, cruzando la frontera, todas sus imagenes $1_A, 1_B$, etc (en donde el subíndice representa la celda a la que pertenece), se mueven cruzando sus respectivas fronteras. La densidad en la celda original (y por lo tanto en todas las celdas y en el sistema completo) se mantiene constante. conservandose la masa.

Aceptando la imposición de periodicidad, basta con almacenar las propiedades de las partículas de la celda original para tener las de todo el sistema.

Si al salir una partícula por un lado entra otra idéntica por el otro, podemos pensar la celda original bidimensional cómo enrollada formando un toro tridimensional. Con esto podemos representar adecuadamente la topología del sistema, aunque la geometría no, pero simplifica pensar en el problema. Si tenemos es una celda tridimensional

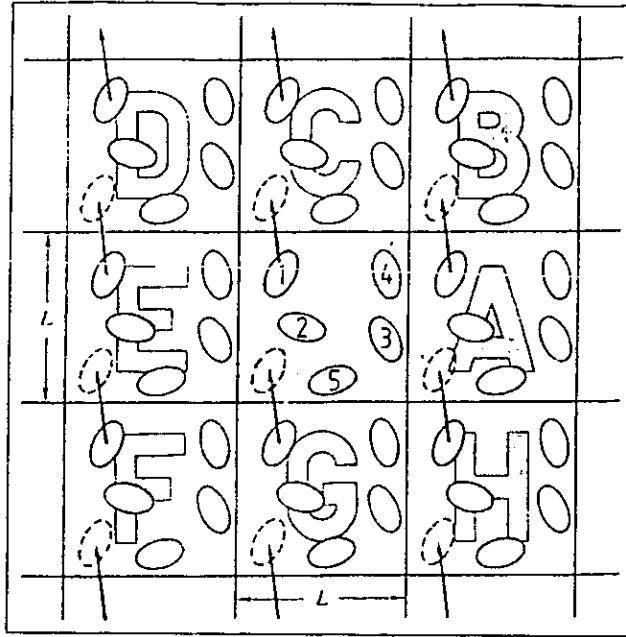


Figura 6-5 Un sistema periódico bidimensional. Los átomos pueden entrar y salir de cada celda por cualquiera de los cuatro lados. [All87]

podemos hacer una analogía con un toro tetradimensional.

Tal y cómo antes nos preguntamos si era posible reproducir con un sistema relativamente pequeño las propiedades de un sistema real, ahora nos preguntamos: ¿son las mismas las propiedades de un sistema periódico infinito y las de un sistema real?. El problema no pertenece nada más a la simulación, sino a la física y es muy interesante. La respuesta no es definitiva: depende tanto del sistema cómo de la propiedad. Por ejemplo, para un fluido en el que los átomos se relacionan por medio del potencial de Lennard-Jones (7) se puede hacer una simulación con celdas cúbicas de lado $L \approx 6\sigma$. Con ese periodo un átomo pierde toda posibilidad de “percibir” la periodicidad del sistema. Por el contrario si el potencial de interacción es de largo alcance (i.e. $V(r) \sim r^{-k}$ en donde k es menor que el número de dimensiones del sistema) habrá una interacción notable entre una partícula de la celda original y sus imágenes, lo que significa una “percepción” de la periodicidad del sistema y de la simetría impuesta a su vez por la simetría de las celdas usadas para teselar el espacio (sección 6.3.4). Allen [All87, pág. 25] menciona que hay trabajos de Mandell de 1976 [Man76] y de Impey, Tildesley y Madden de 1981 [Til81] que muestran que incluso para potenciales de corto alcance las condiciones de frontera periódicas pueden inducir anisotropías en la estructura de fluidos que en realidad son isotrópicos. Hay métodos para investigar los efectos que condiciones de frontera periódicas tienen en las propiedades de equilibrio [Pra81].

Usar condiciones de frontera periódicas evita la aparición de fluctuaciones de período grande. Para teselaciones cúbicas de lado L , no podremos encontrar u observar ondas de densidad con longitud mayor a L , lo que impide simular un líquido cerca de la transición líquido-gas, en donde el tamaño de las fluctuaciones llega a ser macroscópico. Hay autores que muestran ejemplos de cómo transiciones de primer orden pueden tener características de transiciones de ordenes superiores debido a la supresión de las fluctuaciones, cómo por ejemplo el trabajo de Luckhurst y Simpson de 1982 [Luc82] en la transición *nematico* \rightarrow *isotrópico* en cristales líquidos. En sólidos se observan las mismas limitantes al tratar de estudiar fonones con longitud de onda grande en sólidos.

6.3.4 De la simetría de la teselación

Usamos condiciones de frontera periódicas para poder simular sistemas infinitos. Para ello la forma de las celdas debe ser tal que teselen el espacio, y no cualquier figura lo consigue. Lo más común en simulación es utilizar celdas cúbicas, sobre todo por la simplicidad geométrica: pocos vecinos, fáciles de visualizar y ordenar, etc.. Utilizar otras figuras que tengan simetrías iguales a las del sistema a simular puede ser muy conveniente, por ejemplo, para hacer

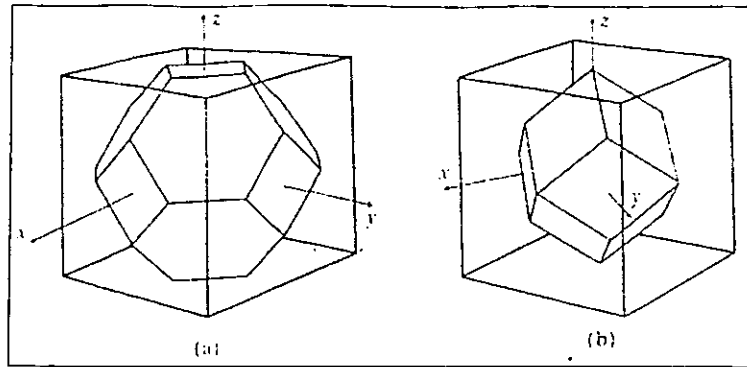


Figura 6-6 Celdas de simulación no cúbicas y los cubos que las contienen. a) Octaedro truncado y b) dodecaedro rómbico. [All87]

simulaciones con cristales hexagonales puede ser mejor utilizar celdas hexagonales que cúbicas.

La simulación de líquidos isotrópicos tiene complicaciones especiales porque la simetría asociada a la isotropía es la de la esfera, y sabemos que no es posible teselar el espacio con esferas, de tal manera que algo hemos de sacrificar. Para inducir un mínimo de efectos en el sistema por medio de forzar una simetría y simular con facilidad, quisieramos: usar una figura lo más parecida a una esfera en cuanto a la cantidad y el tipo de simetrías, una figura simple de tratar geoméricamente, fácil de visualizar, de programar, y de calcular interacciones entre partículas originales e imágenes. Por razones de simplicidad es usual descartar figuras que no sean regulares.

Además del cubo, del resto de los poliedros regulares que teselan el espacio, el dodecaedro (rhombic dodecahedron) y el octaedro truncado también han sido estudiados, y los muestro en la figura 6-6. Son "más esféricos" (mayor relación $\frac{\text{volumen}}{\text{superficie}}$ y más simetrías) que el cubo y además, para una cierta densidad, la distancia entre imágenes es mayor que en el cubo, permitiendo el uso de potenciales de mayor alcance. Para la misma densidad y el mismo radio de corte si una simulación con celdas cúbicas necesita de 256 partículas, entonces un octaedro truncado necesita de 197 y un dodecaedro rhombico solamente 181.

No siempre es posible usar condiciones de frontera periódicas, por ejemplo, puede haber dificultades si existe un potencial externo. Aunque si el potencial exteno es periódico podemos utilizar condiciones con un período igual, pero comunmente este no es el caso y debemos abandonar las condiciones periodicas. También es posible que no queramos utilizar condiciones periodicas en las tres direcciones de los ejes, y que lo hagamos sólo en algunas.

6.3.5 Fronteras de espejo

Hay casos en que en lugar (o además) de usar fronteras periódicas, podemos poner alguna clase de barrera a las partículas que cruzan las fronteras, o inclusive hacerles imposibles el cruce. Un ejemplo es usar fronteras que se comporten como un espejo, en donde el ángulo de incidencia es idéntico al ángulo de reflexión, y el único cambio ocurrido cuando un átomo llega a la frontera es la inversión de una de las velocidades (figura 216-7). Las motivaciones para usar fronteras son varias:

- Simular sistemas libres, en el sentido de que los átomos no tengan interacciones con átomos de celdas imaginarias adyacentes.
- Eliminar la posibilidad de que los átomos escapen a infinito.

Una inversión en algunas componentes de la velocidad no modifica la energía cinética del átomo, ni directamente la energía total, pero dado que el momento lineal cambia, tiene el efecto de un impulso sobre el sistema y representa una interacción del sistema con sus alrededores. Para sistemas de muchos átomos con muchas interacciones con las fronteras y simulados por largo tiempo, esperamos que las choques con todas las caras de la celda esten balanceados y el momento lineal total se conserve.

Estudiando la posibilidad de aplicar este tipo de condiciones de frontera a los sistemas de interés en esta tesis, encontramos diversas propiedades interesantes, por lo que en el capítulo de resultados discutimos detalladamente las simulaciones hechas de esta forma, los resultados y sus implicaciones.

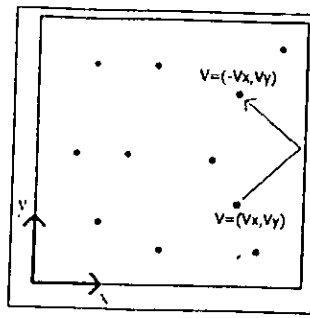


Figura 6-7 Esquema de una frontera de espejo bidimensional. El átomo a la derecha en lugar de salir de la celda invierte su velocidad en esa dirección.

6.3.6 Condiciones de frontera esféricas

Existe una alternativa a las condiciones de frontera periódicas, utilizada sobre todo para tratar líquidos, y es la siguiente: Un sistema bidimensional puede ser imaginado como embebido en la superficie de una esfera, sin necesidad de suponer alguna clase de frontera. Igualmente un sistema tridimensional puede ser embebido en una hiperesfera. La esfera (o hiperesfera) es finita, y aunque es necesario usar geometría no euclidiana para los cálculos (por ejemplo, las distancias entre partículas se miden a lo largo de los círculos máximos que las unen), los efectos de la geometría curva disminuyen cuando el sistema crece, haciéndose tan pequeños que el método parece ser útil para simulaciones de líquidos.

6.3.7 Del cálculo de las interacciones en sistemas con condiciones de frontera periódicas

El corazón de una simulación de Dinámica Molecular o de Monte Carlo es el cálculo de las interacciones entre las partículas. Para utilizar condiciones de frontera periódicas, debemos de tener alguna forma de calcular las interacciones entre una partícula original y todas las demás (originales e imagen).

Volviendo a usar el sistema de la figura 6-5, y suponiendo que el potencial de interacción es un potencial de pares, es necesario considerar no sólo la interacción entre una partícula y las demás $N-1$ de la celda original (figura 6-8), sino la interacción con las N partículas de la celda A, las N de la celda B...

Convención de imagen mínima

Esta suma infinita no necesariamente es fácil de hacer, pero afortunadamente los potenciales de interacción (y sus derivadas por supuesto) tienden a 0 cuando la distancia crece, por lo que es factible despreciar las interacciones con partículas lejanas. Suponiendo a la partícula de interés como al centro de una región que tiene la misma forma y tamaño que las celdas, calculamos su interacción con todas las partículas en esta región, es decir, con las $N-1$ partículas o imágenes de partículas (la que sea más cercana). En la figura 6-8, la partícula 1 interactúa con las partículas 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100. Usar una región con las mismas dimensiones que las celdas garantiza que se ha de interactuar con todas las partículas una y sólo una vez, ya sea directamente o con una de sus imágenes. Este método se conoce como la *convención de imagen mínima*, y presenta algunos problemas: al usar una celda pequeña y eliminar arbitrariamente las interacciones fuera de ella estamos excluyendo del cálculo a muchas partículas con las que la interacción puede no ser completamente despreciable, y aumentar el tamaño de la celda puede llevarnos a realizar cálculos innecesarios.

Radio de corte

Bajo la convención de imagen mínima el número de cálculos necesarios para encontrar todas las interacciones debidas a un potencial de pares es $\frac{1}{2}N(N-1)$, lo cual puede ser excesivamente grande (y tardado) incluso para sistemas pequeños. Una segunda aproximación para el cálculo de las interacciones simplifica aún más el problema: Si las principales contribuciones a la energía y las fuerzas (en el caso de una simulación de Dinámica Molecular) provienen de las partículas vecinas más próximas, podemos aplicar lo que se conoce como un *radio de corte esférico* r_c . Despreciaremos cualquier interacción entre partículas a una distancia mayor que r_c (lo que significa suponer que el

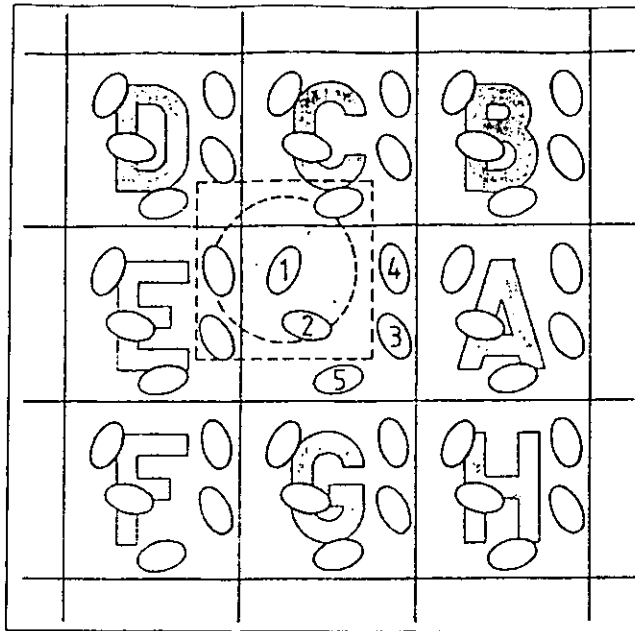


Figura 6-8 La convención de imagen mínima bidimensional. La caja central contiene 5 átomos. Se construye usando el átomo 1 como centro y trazando un cubo de lado L y un círculo de radio igual al de corte. [All87]

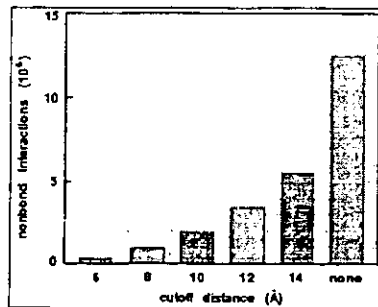


Figura 6-9 Número de interacciones como función del radio de corte [Msi97] en un sistema de 5000 átomos. El tiempo requerido para evaluar la energía es aproximadamente proporcional al número de interacciones. [Msi97]

potencial es ya prácticamente cero si $r > r_c$). En la figura 6-8. El círculo punteado representa el radio de corte. En este caso las partículas 2 y 4_E, se consideran en interacción con la 1, mientras que la 3_E y la 5_C no. La cantidad de partículas con las que ya no es necesario tomar en cuenta la interacción se reduce aproximadamente por un factor de $\frac{4}{3}\pi(\frac{r_c}{L})^3$ y si $\frac{r_c}{L}$ es pequeño el ahorro en cálculos puede ser notable.

El radio de corte debe ser menor a $\frac{1}{2}L$ para ser consistente con la convención de imagen mínima en el caso de celdas cúbicas, pero para los casos de la figura 6-6 puede ser un poco mayor.

Brooks y otros muestran en su artículo [Bro85] los efectos en la energía y la configuración de aplicar un radio de corte. La figura 6-9 muestra la cantidad de interacciones de pares (nonbond interactions) esperadas para un sistema de 5000 átomos como función del radio de corte. En la sección 6.12.1 hago una discusión del efecto de los radios de corte en la energía potencial.

Convención de imagen explícita

Existe un procedimiento para el tratamiento de fronteras periódicas. es la *convención de imagen explícita*. Consiste en generar imágenes periódicas explícitas de los átomos de la celda original llamadas *átomos fantasma*. Estos átomos fantasma son replicados a la distancia que sea necesario para satisfacer el radio de corte del potencial elegido. Sólo los grupos de átomos fantasma dentro del alcance del potencial son utilizados para calcular las interacciones.

Los átomos fantasma tienen sus contrapartes simétricas. Por ejemplo, el lado izquierdo de la figura 6-10 muestra la molécula (que toma el lugar de un átomo) A1 interactuando con varias imágenes de B (B1, B2, B3 y B5) dentro

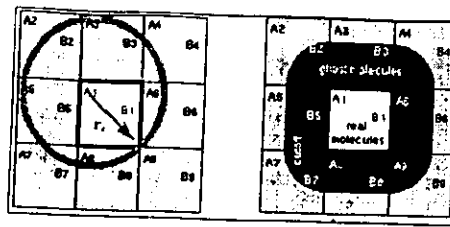


Figura 6-10 Convención de imagen explícita que muestra cómo la distancia de corte determina cuales átomos de las celdas imagen adyacentes son elegidas como fantasmas (notese que los radios de corte son diferentes en las figuras de la izquierda y la derecha). Izquierda: modelo de imagen explícita donde se ve que es posible un radio de corte mayor que con la convención de imagen mínima. Derecha: la región sombreada muestra cuáles átomos son seleccionados cómo fantasmas dentro del radio de corte de cualquier átomo en la celda original. [Msi97]

del radio de corte especificado (denotado por un círculo centrado en A1): A1 interactúa con varias de sus propias imágenes: A3, A5, A6 y A8. Sin embargo, al momento de mover los átomos, sólo los reales (A1 y B1) son movidos, según la fuerza sobre cada uno. Las posiciones de los átomos fantasma son simplemente actualizadas al aplicar las relaciones de simetría a las nuevas posiciones de las partículas.

El lado derecho de la figura 6-10 muestra cuales átomos en las celdas unidad adyacentes se convierten en átomos fantasma explícitos para un cierto radio de corte. No todos los átomos en una celda adyacente se convierten en un fantasma. Sin embargo, si el radio de corte es mayor que la longitud de la celda usada, es posible usar fantasmas de celdas mas allá. Conforme los átomos entran y salen de las fronteras, los átomos fantasmas cambian, por lo tanto, la lista de fantasmas se actualiza periódicamente. Notese que las interacciones entre fantasmas no tienen que ser calculadas porque simplemente agregan una energía de base.

Correcciones de largo alcance

Dejar de incluir algunas interacciones por un radio de corte introduce una perturbación y las propiedades calculadas pueden ser diferentes de aquellas que se obtienen sin el corte; si el potencial y su derivada son muy cercanos a cero a la distancia de corte las perturbaciones son muy pequeñas y puede que no alteren la esencia del resultado. De cualquier manera es un punto a tener en cuenta al usar radios de corte. Una manera de corregir algunos de los errores introducidos de esta forma es agregar correcciones de largo alcance. Estas por lo general se hacen despreciando la estructura específica, usando funciones de la densidad: por ejemplo, si el potencial es de tipo Lennard-Jones la corrección para el cálculo de la energía potencial es

$$E_{LRC}^* = \frac{8}{9}\pi N\rho^* r_c^{*-9} - \frac{8}{3}\pi N\rho^* r_c^{*-4} \quad (6.47)$$

en donde LRC significa corrección de largo alcance (por las siglas en inglés de *Long Range Correction*) y las * se refieren a que las unidades son unidades reducidas de Lennard-Jones (sección 6.1).

En la sección 7.8 abundo en el tratamiento de las fuerzas de largo alcance, al discutir sobre otras dificultades al simular y la manera de tratar con ellas.

6.3.8 Geometrías con superficies

Los propósitos de las condiciones de frontera periódicas son eliminar los efectos de superficie y permitir simular sistemas infinitos. Sin embargo, puede que *queramos* tener superficies. Para hacer simulaciones con superficies, el modelo usualmente adoptado es el de la *rebanada* (*slab*): una gruesa rebanada de material delimitada por dos superficies. Esta se obtiene fácilmente removiendo las condiciones de frontera periódicas a lo largo de una dirección. La rebanada puede pensarse como infinita sobre un plano, pero finita sobre el eje ortogonal.

Si la rebanada es suficientemente gruesa, su interior se ha de comportar en forma muy similar a un material infinito. Las dos superficies (superior e inferior) pueden entonces considerarse cómo desacopladas, es decir independientes. Bajo esta suposición el comportamiento del sistema será cercano al de una superficie y más cercano a un experimento.

Hay circunstancias en que al simular puede ser útil "congelar" algunas capas del material de un lado de la rebanada, construyendo esos átomos a quedarse en algunas posiciones específicas. Esto se hace cuando se cree que

una de las superficies puede tener efectos sobre la otra mayores que los que tendría un sistema fijo. Preferimos este tipo de simulaciones cuando ocurren reajustes importantes en la estructura de la superficie o fusión. Si se elige “congelar” una de las superficies se debe tener cuidado [Erc97, pág. 16] de congelar un número de capas al menos de grosor r_c , para garantizar que ningún átomo móvil podrá percibir la superficie (la falta de continuidad).

Es posible, incluso, utilizar condiciones de frontera periódicas solamente sobre una dirección, obteniendo la geometría de un *cable* con el eje en la dirección de la periodicidad.

Eliminar por completo las condiciones de frontera periódicas permite simular sistemas de racimos (*clusters*) de átomos. Los racimos constituyen líneas de investigación muy importantes y han sido estudiados desde los primeros días de la simulación.

6.3.9 Fronteras fijas y estocásticas

Hay dos variantes más de frontera que se usan con frecuencia. Son necesarias para evitar las periodicidades artificiales inducidas por condiciones de frontera periódicas. En ambos casos el espacio se divide en tres regiones:

1. La celda de simulación central. Los átomos en ella se mueven libres, y es la región en donde podremos hacer mediciones.
2. Una capa alrededor de la celda central. Los átomos en ella pueden estar fijos o constreñidos por potenciales armónicos. Este tipo de fronteras se conoce como *fijas*. En la variante estocástica los átomos de esta región son sujetos a fuerzas estocásticas, que simulan una interacción con el exterior. Por lo general esta región tiene un grosor mayor al alcance del potencial, para reducir los efectos de la frontera en los átomos de la celda central.
3. El espacio exterior. Hay dos variantes importantes en la manera de tratar este espacio exterior. La primera supone un espacio exterior vacío. La segunda lo supone compuesto de un continuo con ciertas propiedades, por ejemplo una constante dieléctrica. Suponer un continuo en el exterior tiene la ventaja de permitir una capa intermedia más delgada y de simular mejor los efectos globales del material en la celda central.

6.4 Hacia el equilibrio

“La liberación de la energía del átomo no crea un nuevo problema. Simplemente hace más urgente la necesidad de solucionar uno existente.” Albert Einstein.

En algunos momentos durante una simulación, el sistema se encuentra fuera de equilibrio termodinámico, donde los indicadores del estado del sistema no se encuentran estables, sino relajándose hacia un nuevo valor, pasando de un estado a otro.

El cambio en el estado del sistema puede ser inducido por nosotros o espontáneo. Lo inducimos nosotros cuando cambiamos algún parámetro de la simulación, como la temperatura o la densidad, perturbando al sistema. Es espontáneo cuando, por ejemplo, el sistema pasa por una transición de fase, pasando de un estado de equilibrio a otro.

Lo usual es que queramos tener al sistema en un estado de equilibrio antes de comenzar a hacer mediciones. La aproximación al equilibrio de una cantidad física A acostumbra aproximarse con una función exponencial del tiempo

$$A(t) = A_0 + C e^{-t/\tau} \quad (6.48)$$

en donde $A(t)$ denota la cantidad física promediada sobre un tiempo corto, para eliminar el efecto de las fluctuaciones pero no de la deriva en tiempos largos. La variable τ recibe el nombre de *tiempo de relajación* y da sentido a la escala temporal en que ocurre el proceso. Hay casos en que τ es del orden de cientos de pasos de integración, permitiéndonos observar la convergencia de $A(t)$ a A_0 y hacer medidas directas del valor de equilibrio. También es posible que τ sea muy grande comparada con la escala temporal de nuestra simulación; un valor realista de este tipo sería una $\tau = 1$ s, muchos órdenes de magnitud mayor que la escala típica de una simulación, por lo que en este caso no observamos ninguna relajación y la propiedad parece equilibrada. En situaciones intermedias puede que observemos una ligera

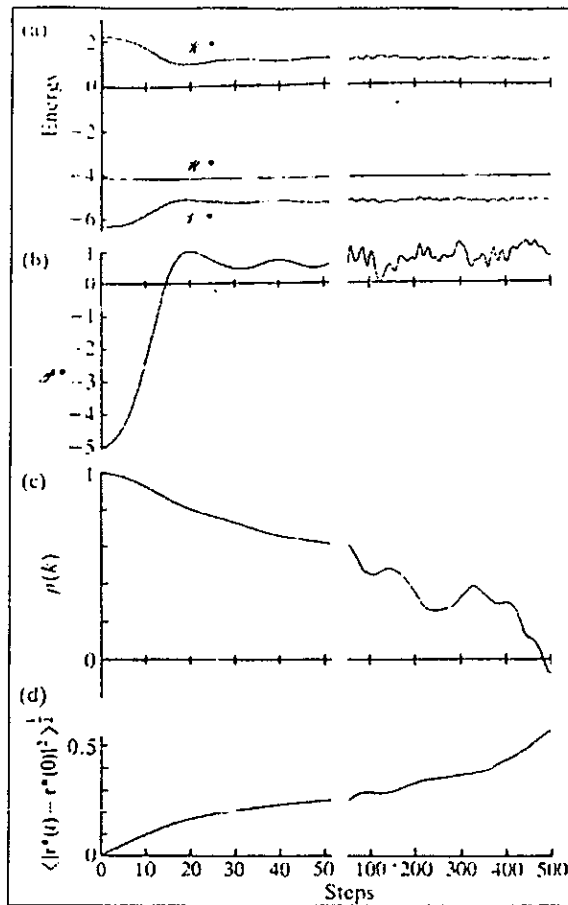


Figura 6-11 Fase de equilibrio de una simulación de dinámica molecular. Los primeros 50 pasos se muestran en detalle. El sistema consiste de 108 átomos interactuando vía el potencial de Lennard-Jones y comenzando de una red FCC con velocidades distribuidas según la función de Maxwell-Boltzman. a) Energías potencial V , cinética K y total H . b) Presión instantánea. c) Parámetro de orden traslacional. d) Desplazamiento RMS (Root Mean Square). [Al187]

deriva pero no podemos simular lo suficiente como para esperar la completa convergencia de $A(t)$ a A_0 . En estos casos se usan los datos conocidos y se hace una estimación de A_0 utilizando la ecuación 6.48.

El problema de la escala temporal en la identificación del equilibrio es muy importante, y es un factor clave al encontrar estados de equilibrio usando dinámica molecular. Parte del problema es que el equilibrio es un concepto macroscópico [Hai92, pág. 206]: “se aplica en un tiempo finito a un sistema de tamaño macroscópico”. Tomar muestras del sistema en un instante no es suficiente para identificar equilibrio.

Un ejemplo, de la vida diaria, de las diferencias en la escala temporal es la navaja de un cuchillo de carbón-acero en un ambiente húmedo. Por unas horas el cuchillo aparenta estar en equilibrio con sus alrededores, y aparentemente no hay nada que muestre que se encuentra fuera de un equilibrio químico o mecánico. Sin embargo, después de estar expuesto varios días al aire húmedo la navaja se ha oxidado. ¿Esta el cuchillo en equilibrio con sus alrededores o no?, la respuesta es simple: depende de la escala temporal que nos interese. (De hecho la navaja del cuchillo se encuentra en estados metaestables que llevan tarde o temprano a la rotura de la misma).

Permitir que un sistema llegue al equilibrio es importante, sobre todo al inicio de la simulación, en donde con la configuración inicial podemos haber colocado al sistema en un estado para el que los cálculos de algunas propiedades no necesariamente tienen sentido. Permite borrar toda memoria de la configuración original y dar mayor legitimidad a los cálculos.

Un ejemplo de evolución al equilibrio son los diagramas de la figura 6-11 en donde se observan la energía y la presión desde sus valores iniciales hasta sus valores finales de equilibrio. En este caso decimos que se ha llegado a equilibrio porque ya no se puede observar la deriva sistemática y las oscilaciones son alrededor de una media constante.

El equilibrio es particularmente importante si nos interesa estudiar líquidos comenzando con configuraciones cristalinas (sección 5.1), donde es necesario observar la fusión del sistema por medio de varios parámetros. Par-

ticamente útil en este sentido es el factor de estructura (secciones 6.9 y 6.10.3) reducido al parámetro de orden traslacional para lattices, mostrado en la figura 6-11c y que es más sensible a la estructura periódica que la energía y la presión. En simulaciones moleculares es común utilizar también el parámetro de orden traslacional introducido por Vieillard-Baron en 1972 [Vie72] dado por

$$P_1 = \frac{1}{N} \sum_{i=1}^N P_1(\cos \gamma_i) = \frac{1}{N} \sum_{i=1}^N \cos \gamma_i \quad (6.49)$$

en donde γ_i es el ángulo entre el eje de la molécula i y el eje original del cristal perfecto.

Para casos en que se esperan transiciones de fase, es recomendable observar la evolución de la desviación cuadrática media (sección 6.6.11) de los átomos alrededor de sus posiciones cristalinas. Este valor debe aumentar en una simulación de líquidos pero mantenerse constante en un sólido. La regla común es que cuando la RMS (root-mean-squared o raíz del cuadrado de la media) del desplazamiento por partícula sobrepasa 0.5σ y se está incrementando el sistema se ha "fundido". Hay que tener cuidado con los efectos de las condiciones de frontera periódicas al calcular esta cantidad.

Un peligro adicional durante el periodo de equilibrio es que el sistema entre en un estado metaestable, por ejemplo una región de coexistencia de líquido y gas. Si lo que nos interesaba era estudiar al sistema en su estado líquido nos podemos topar con resultados sorprendentes. Allen [All87, pág. 173] sugiere utilizar la función de distribución radial (sección 6.8) que en la región de interfase se comporta en forma peculiar.

Usualmente sistemas con mayor temperatura (y por lo tanto mayor movilidad de sus átomos) llegan a los estados de equilibrio con mayor rapidez, por lo que también es práctica común elevar la temperatura hasta llevar un sistema al equilibrio, y luego buscar el estado deseado mediante enfriamiento lento.

Algunas condiciones necesarias para poder afirmar que un sistema aislado se encuentra en equilibrio (para un ensamble microcanónico) son:

1. La cantidad total N de átomos y la energía \mathcal{E} deben ser constantes independientes del tiempo. Dado que \mathcal{E} es constante en el tiempo las fluctuaciones de la energía cinética y potencial deben ser de magnitudes iguales pero fuera de fase.
2. Cada componente cartesiana de la velocidad debe cumplir, en promedio temporal, con una distribución Gaussiana (ecuación 5.6).
3. Propiedades termodinámicas como la temperatura, energía interna (de configuración), y presión deben fluctuar alrededor de un promedio estable. Este promedio debe ser independiente de cómo llegamos a el estado de equilibrio, es decir, deben ser reproducibles. El tamaño de las fluctuaciones depende del tamaño del sistema, es decir, del número N de átomos y disminuyen como \sqrt{N} . Por ejemplo, aumentar el tamaño de un sistema de 108 átomos a 500 disminuye las fluctuaciones a la mitad (ver figura 6-12).
4. Los promedios de las propiedades deben ser estables bajo pequeñas perturbaciones. Si el estado es modificado levemente, por ejemplo agregando y luego retirando una cierta cantidad de calor, las propiedades termodinámicas deben recobrar sus valores de equilibrio. La intención de esta condición es la de alejar al sistema de estados metaestables, que de otra manera muestran las propiedades de estados de equilibrio. Para generar la figura 6-13, se utilizaron 108 átomos que interactuaban vía el potencial de Lennard-Jones. 5000 pasos después del equilibrio se escalaron las velocidades por un factor de 1.5 (ver sección 6.12). Se mantuvo el rescalamiento por 20 pasos (0.17psec) después de lo cual las velocidades fueron reescaladas por otros 20 pasos para recuperar la energía total anterior ($E = -1.48N\epsilon$). La figura 6-13 muestra que tanto la energía interna como la presión instantáneas no se ven afectadas por la perturbación.
5. Si el sistema se divide en sus partes macroscópicas, los promedios temporales de cada propiedad deben ser los mismos en cada parte. Por ejemplo, si arbitrariamente dividimos el sistema a la mitad y luego promediamos en el tiempo, el número de átomos, la temperatura, la energía total y la presión deben ser las mismas en ambas mitades.

Ninguna de estas cinco condiciones sola, ni todas ellas en conjunto pueden garantizar la existencia de un estado de equilibrio, sin embargo constituyen fuertes argumentos en su favor, al menos para la escala temporal usual de una simulación.

Lo normal no es aplicar las pruebas de equilibrio a todas las simulaciones, sino que se realizan algunas pruebas piloto, aunque se debe tener en cuenta que si se hacen cambios sustanciales a las condiciones es necesario verificar que se cumplan nuevamente.

No hay que olvidar un último problema en potencia que podemos encontrar: haber colocado al sistema en una parte irrelevante del espacio fase. La manera de tratarlo es comenzar con diferentes condiciones iniciales.

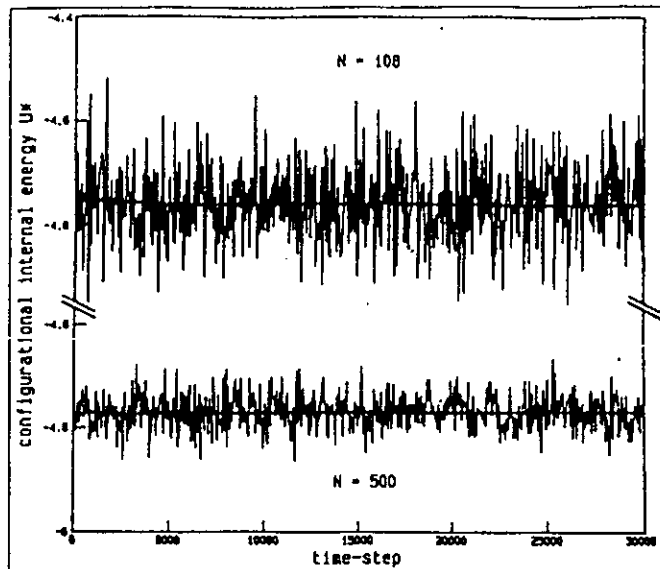


Figura 6-12 Diferencia en las fluctuaciones de propiedades al aumentar el tamaño del sistema. Aquí se muestra la energía total de dos simulaciones de sistemas en el mismo estado termodinámico, pero con diferente número de átomos: arriba 108 átomos y abajo 500 átomos. En ambos la línea horizontal es el promedio móvil de los valores fluctuantes. Los valores se grafican cada 50 pasos con un $dt = 0.001$. [Hai92]

6.5 Mirando los átomos

“Pensar es más interesante que saber, pero menos interesante que mirar”, Goethe [Lod99].

La manera más simple (al menos desde el punto de vista conceptual) de examinar un sistema para entender lo que le ocurre es “mirarlo”. Podemos asignarle un radio a cada uno de los átomos (por ejemplo la cuarta parte de la distancia de equilibrio, o la mitad de la distancia a la que el potencial de pares se vuelve positivo), representarlos como pelotas, elegir un programa y construir alguna imagen del sistema desde algún punto de vista interesante. La imagen resultante la podemos observar en el momento o con calma después.

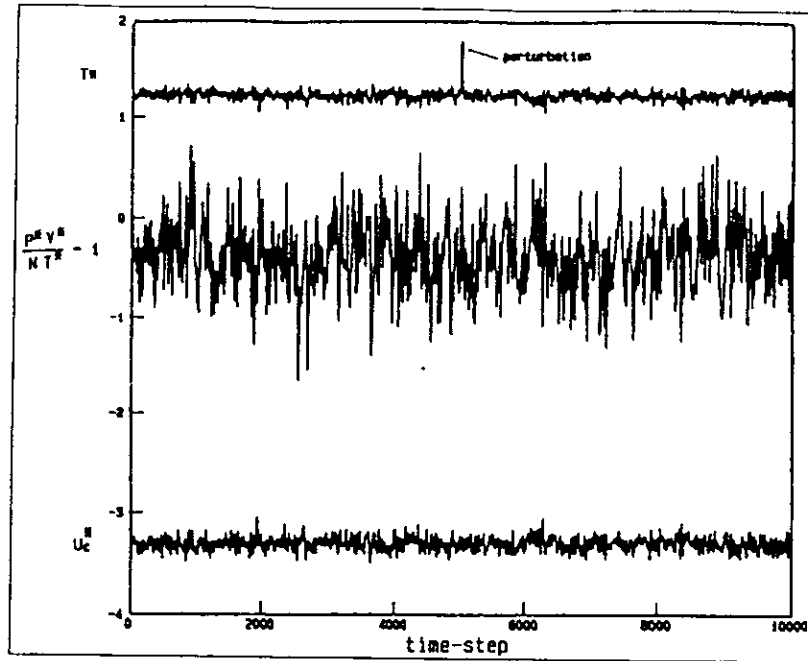
En algunas situaciones la vista es la mejor manera de inspeccionar un sistema, dadas las dificultades de analizar todos y cada uno de los aspectos de una configuración. Por ejemplo, al analizar un arreglo de átomos podemos distinguir a simple vista patrones difíciles de distinguir matemáticamente.

Estamos muy acostumbrados a mirar las cosas, y aunque la vista nos puede engañar, nunca deberíamos dejar de observar las cosas. Las simulaciones de dinámica molecular se prestan más que las de monte carlo para ser visualizadas, porque las configuraciones están ordenadas temporalmente y en monte carlo no hay tal orden. Desde el punto de vista educativo esto representa una virtud más de la Dinámica molecular sobre Monte carlo.

Muchos de los programas que generan imágenes o películas de configuraciones permiten mapear en las imágenes propiedades atómicas, por ejemplo la energía de los átomos mostrada como color, o el tipo de átomo distinguible en su diámetro, etc. Este tipo de mapeos permite observar fácilmente la distribución espacial de algunas propiedades. En el caso de películas podemos observar la evolución temporal de las distribuciones espaciales de las propiedades.

La figura 6-14 muestra las trayectorias de un sistema en diferentes estados. Sin necesidad de calcular propiedades, es clara una diferencia entre las fases sólida y líquida.

Sin embargo, cantidades precisas no son obtenibles de observaciones, por lo que es recomendable combinar la observación con cálculos. Comparese con la figura 6-16 en donde se simulan conjuntamente sólido y líquido, pudiéndose distinguir entre ambas fases mediante las trayectorias.



ESTA TESIS NO DEBE
 SALIR DE LA BIBLIOTECA

Figura 6-13 Un sistema ya en equilibrio es estable a pequeñas perturbaciones. Aquí se muestran la presión y energía potencial instantáneas. La temperatura se aumenta momentáneamente por un factor de 1.5. Hasta arriba se muestra la temperatura, en medio la presión y hasta abajo la energía potencial. [Hai92]

6.6 Midiendo el ensamble microcanónico

“No todo lo contable cuenta y no todo lo que cuenta es contable”, Albert Einstein [Ali99].

Las propiedades físicas son definidas como funciones de las coordenadas y velocidades de las partículas, de tal manera que uno puede definir, por ejemplo, el valor instantáneo de A al tiempo t como

$$A(t) = f(\mathbf{r}_1(t), \dots, \mathbf{r}_N(t), \mathbf{v}_1(t), \dots, \mathbf{v}_N(t)) \quad (6.50)$$

si asumimos que la energía se conserva el promedio sobre la trayectoria sera

$$\bar{A} = \lim_{t' \rightarrow \infty} (t' - t_0)^{-1} \int_{t_0}^{t'} A(\mathbf{r}^N(t), \mathbf{p}^N(t)) dt, \quad (6.51)$$

y (por la hipótesis ergódica) es igual a un promedio en el ensamble microcanónico (sección 4.3)

$$\bar{A} = \langle A \rangle_{NVE} \quad (6.52)$$

que en sistemas discretos se convierte en

$$\bar{A} = \frac{1}{N_T} \sum_{t=1}^{N_T} A(t). \quad (6.53)$$

La notación es la usual: denotar los promedios sobre el ensamble con $\langle \cdot \rangle$ y promedios sobre la trayectoria con $\bar{\cdot}$

Los pasos de tiempo involucrados en los promedios no tienen que ser todos los calculados por la simulación, pueden ser algún subconjunto de configuraciones, por ejemplo para ahorrar tiempo de cálculo, simplificar el problema, o para hacer mediciones parciales.

Dependiendo del interés que tengamos en una propiedad y de la importancia de la velocidad de ejecución, los cálculos se pueden hacer en dos momentos diferentes:

1. En tiempo de ejecución. $A(t)$ es calculada cada paso de tiempo por el programa de dinámica molecular. La suma $\sum_t A(t)$ también se actualiza a cada paso. En ciertos momentos del programa el promedio se obtiene

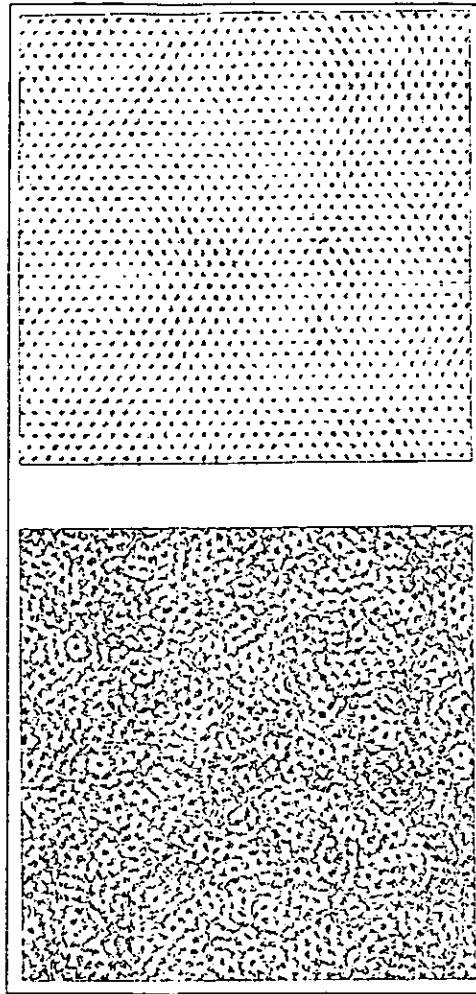


Figura 6-14 Imágenes de las trayectorias atómicas a dos densidades $\rho = 0.95$ y $\rho = 0.75$, enfatizando la diferencia entre las fases sólida y líquida cómo localizada y difusiva. [Rap95] Podemos comparar con la figura 6-16, en donde se simulan conjuntamente sólido y líquido.

dividiendo el valor acumulado entre el número de pasos. Este es el método preferido cuando la cantidad es simple de calcular y/o particularmente importante, cómo la temperatura. Una virtud de este método es que permite aprovechar las cantidades calculadas durante la simulación para dirigirla y explorar el espacio de configuraciones de una manera inteligente.

2. Posteriormente. Posiciones (y posiblemente velocidades) son periódicamente guardados a un archivo de configuraciones durante la ejecución. Otro programa que se ejecuta después del de simulación, procesa las configuraciones y calcula las cantidades deseadas. Para un sistema grande este método puede requerir demasiado espacio en disco. Es conveniente cuando nos interesa calcular algunas cantidades complejas, que podrían retrasar mucho la simulación. También pueden ser dependientes de otros parámetros que no sean evidentes o conocibles al ejecutar la simulación: por ejemplo, para calcular un patrón de difracción de electrones se requiere mucho tiempo y ajustar varios parámetros, por lo que para estudiar alguna configuración en particular se acostumbra guardar en disco para luego procesarla con calma hasta obtener el patrón adecuado.

Generar imágenes es tan tardado que normalmente no se realiza durante una simulación, y se reserva el poder de cálculo para lo esencial. Lo normal es almacenar algunas configuraciones interesantes y después generar las imágenes.

6.6.1 Energía potencial

La energía potencial \mathcal{V} promedio se obtiene de los valores instantáneos, que se obtienen en forma directa al mismo tiempo que se calculan las fuerzas. Por ejemplo, en el caso de interacciones por un potencial de pares

$$\mathcal{V}(t) = \sum_{j>i} \phi(r_{ij}(t)), \quad (6.54)$$

en donde $r_{ij}(t)$ ya habrá sido encontrado por ser necesario para las fuerzas.

En las secciones 6.3.7, 6.8, 7.7 y 7.8 discuto los métodos para aplicar un radio de corte en el alcance del potencial y ahorrar tiempo en los cálculos de energías y fuerzas. Si no tenemos cuidado en una simulación podríamos usar hasta el 99% del tiempo de ejecución, en el cálculo de las energías y fuerzas [Hee90, pág. 22].

Un radio de corte puede introducir una singularidad en las funciones de fuerza a la distancia de corte por la discontinuidad en el potencial y la no existencia de una derivada en ese punto. Si el potencial es cero a la distancia de corte la discontinuidad en la energía desaparece, aunque es necesario que el potencial además llegue a cero suavemente a esa distancia para remover también la discontinuidad en sus derivadas. Es común asegurarse de esta convergencia para evitar problemas, ya sea eligiendo un radio de corte mayor o ajustando el potencial. Los efectos de haber truncado el potencial pueden ser considerables [Pen79] [Hee82]. Para hacer una estimación del efecto aprovechamos que la energía potencial puede ser calculada por

$$\mathcal{V} = 2\pi\rho N \int_0^\infty \phi(r)g(r)r^2 dr, \quad (6.55)$$

en donde $g(r)$ es la función de correlación de pares que mide la correlación entre las partículas (independiente del tiempo), y que explico en la sección 6.8.

La *corrección de cola* del potencial, es una aproximación de la energía potencial más allá del radio de corte, definida como

$$\mathcal{V}_c = 2\pi\rho \int_{r_c}^\infty \phi(r)g(r)r^2 dr. \quad (6.56)$$

No siempre se calcula $g(r)$ durante la simulación, en cuyo caso para poder aproximar la corrección de cola podemos asumir $g(r) = 1$. El error asociado disminuye si el radio de corte es grande. Un ejemplo del efecto de aplicar corrección de cola se puede observar en la tabla 8.1. de la sección 6.12 en donde nota que los efectos son relativamente importantes (9.8% y 2.7% respectivamente).

6.6.2 Energía cinética

La energía cinética instantánea esta dada por

$$\mathcal{K}(t) = \frac{1}{2} \sum_i m_i (v_i(t))^2 \quad (6.57)$$

En el ensamble microcanónico esta no se conserva, sino que se intercambia con la energía potencial. El promedio de este valor según la ecuación 6.53 es

$$\bar{\mathcal{K}} = \frac{1}{N_T} \sum_{t=1}^{N_T} \mathcal{K}(t). \quad (6.58)$$

La energía cinética es útil para verificar la conservación de la energía total, calcular la temperatura (sección 6.6.7), buscar mínimos de energía y más. Es muy raro encontrar simulaciones de dinámica molecular en que no se calcule la energía cinética, inaccesible para simulaciones de tipo Monte Carlo.

6.6.3 Energía total

La energía total $\mathcal{E} = \mathcal{K} + \mathcal{V}$ es una cantidad conservada en la dinámica newtoniana (es decir $\mathcal{E} = \bar{\mathcal{E}}$), sin embargo, es común calcularla tan seguido como sea posible para verificar el comportamiento del programa mediante la medición

de las fluctuaciones en la energía total. En [Ber86] Berendsen y Gunsteren hacen un análisis detallado de las fluctuaciones de la energía usando diferentes algoritmos de integración. Al aumentar el número de átomos del sistema las fluctuaciones disminuyen como $\sqrt[3]{N}$, como se observa en la figura 6-12.

Lentas derivas de la energía total se observan si la simulación es suficientemente larga. Son controlables modificando la magnitud de dt . Las derivas pueden ser más peligrosas que las fluctuaciones, porque el estado termodinámico del sistema cambia al cambiar la energía, alejándonos de los estados que nos interesa estudiar y haciendo que los promedios temporales que calculamos pierdan sentido. Es posible corregir algunas de las derivas en simulaciones largas partiendo estas en simulaciones más pequeñas, ajustando la energía total entre las diferentes etapas.

Ajustar las energías potenciales no es fácil, y lo normal es ajustar las energías cinéticas por medio de un escalamiento de las velocidades, modificando la temperatura (ver sección 6.6.7 del cálculo de la temperatura y 6.12 del control de la misma).

6.6.4 Momento lineal

Para calcular el momento lineal basta con las masas y las velocidades. Usualmente suponemos que las masas son todas 1 o que al menos son constantes (sección 6.1), y las velocidades ya las hemos calculado para hacer la evolución temporal del sistema.

Si el sistema se encuentra aislado, es de esperarse que se mantenga constante el momento lineal, lo que lo hace útil para verificar el funcionamiento del programa.

Utilizar algunas condiciones de frontera en particular puede llevarnos a modificar el momento lineal del sistema, por ejemplo, al utilizar condiciones de frontera de espejo estaríamos invirtiendo la dirección de las partículas que salen de la celda de simulación. Esta variación en la velocidad significa una variación en el momento, aunque no una variación en la energía cinética del sistema porque esta depende de la magnitud de las velocidades que no es modificada.

Aún cuando podemos simular un sistema con cualquier momento lineal inicial, no ganamos nada si este es diferente de cero, por lo que usualmente las velocidades iniciales se eligen (o se modifican) en forma adecuada para hacer el momento lineal total inicial igual a cero. En la sección 5.1.2 escribo más al respecto de la asignación de velocidades iniciales.

6.6.5 Momento angular

A cada cantidad conservada en la física se le asocia una simetría. En el caso del momento angular es la de rotación. Esto quiere decir que el sistema debe de ser equivalente sin importar la dirección en la que se encuentre. Si nos interesa simular un sistema en el que se conserve el momento angular, necesitamos que el universo simulado tenga simetría esférica.

El momento angular, al igual que el momento lineal, es una cantidad importante, aunque por lo general presenta más problemas.

Uno de ellos es que si utilizamos condiciones periódicas para simular un sistema infinito, necesitaremos celdas que tengan una forma que pueda teselar el espacio. Desgraciadamente no existen figuras que pudiendo teselar el espacio tengan simetría esférica. Esto significa que si queremos simular sistemas periódicos necesitamos sacrificar algo de simetría y con ella la conservación del momento angular.

También se acostumbra hacer el momento angular total inicial cero para disminuir la deriva rotacional del sistema, y aunque la falta de simetría impide la conservación, el momento angular no tiene porque dispararse. Es conveniente calcularlo para verificar que no se aleje mucho del cero y en ocasiones puede ser nuevamente ajustado.

6.6.6 Función H de Boltzmann

Un sistema tiene, por supuesto, una cierta distribución de velocidades, y en ella se puede encontrar información útil para entenderlo. Es más fácil observar la evolución de un sólo número que la de una distribución. Para ello utilizamos

la función H de Boltzmann.

Para calcular la parte cinética de la función H primero necesitamos calcular la distribución de velocidades. La distribución de velocidades mide la fracción del total de átomos que tiene una velocidad en el intervalo $\mathbf{v} \pm \frac{d\mathbf{v}}{2}$,

$$f(\mathbf{v})d\mathbf{v} = \frac{1}{N}N(\mathbf{v})d\mathbf{v}. \quad (6.59)$$

Para una componente cartesiana, por ejemplo v_x , la ecuación anterior se vuelve

$$f(v_x)dv_x = \frac{1}{N}N(v_x)dv_x \quad (6.60)$$

Podemos determinar $f(v_x)$ durante una simulación eligiendo un intervalo finito Δv_x y contar el número de átomos con velocidades en $v_x \pm \frac{\Delta v_x}{2}$. Se define [Hai92] la parte cinética de la función H instantánea cómo

$$H_x = \sum_{\Delta v_x} f(v_x) \ln f(v_x) \Delta v_x. \quad (6.61)$$

Completando con las componentes y y z obtenemos

$$H = \frac{1}{3}(H_x + H_y + H_z). \quad (6.62)$$

En un sistema en equilibrio hemos de encontrar que la distribución de velocidades obedece la ley de Maxwell-Boltzmann [Cha61]

$$N(\mathbf{p}) = \int N(\mathbf{r}, \mathbf{p}) d\mathbf{r} = \frac{N}{C} e^{-\frac{\mathbf{p}^2}{2mk_B T}} \quad (6.63)$$

en donde k_B es la constante de Boltzmann y C es una constante que asegura satisfacer la normalización

$$\int N(\mathbf{r}, \mathbf{p}) d\mathbf{r} d\mathbf{p} = N. \quad (6.64)$$

Asociada a una distribución de Maxwell-Boltzmann de las velocidades de un sistema en particular hay un valor de la función H de Boltzmann. Una buena medida para saber si nos encontramos en un estado de equilibrio es comparar este valor con el valor instantáneo de la función H .

Es posible calcular promedios de la función H , ¿debemos entonces usar el valor instantáneo o el valor promedio?. Según Haile [Hai92, pág. 121] el valor promedio puede dar información equivocada, por lo que recomienda usar el valor instantáneo.

6.6.7 Temperatura

Estamos interesados en calcular observables en el límite termodinámico, en donde todos los ensambles son equivalentes. Una consecuencia de la equivalencia de los ensambles es que podemos calcular las propiedades termodinámicas básicas de un modelo como promedios en cualquiera de los ensambles. Según esto no es necesario especificar el ensamble al que nos referimos cuando describamos la forma en que calculamos cantidades termodinámicas.

La temperatura y la presión pueden calcularse utilizando el teorema del virial, que según la forma de Münster 1969 [Mün69] es

$$\left\langle p_k \frac{\partial \mathcal{H}}{\partial p_k} \right\rangle = k_B T \quad (6.65)$$

$$\left\langle q_k \frac{\partial \mathcal{H}}{\partial q_k} \right\rangle = k_B T \quad (6.66)$$

para las coordenadas generalizadas q_k o momentos p_k . Estas ecuaciones son ejemplos de la forma general

$$\left\langle A \frac{\partial \mathcal{H}}{\partial q_k} \right\rangle = k_B T \left\langle \frac{\partial A}{\partial q_k} \right\rangle \quad (6.67)$$

para el ensamble canónico [All87, pág. 46], y son válidas hasta un orden de N^{-1} , por lo que para la mayoría de los problemas de simulación son suficientes.

La ecuación 6.65 es particularmente simple si los momentos aparecen elevados al cuadrado en el Hamiltoniano, en cuyo caso podemos sumar $3N$ términos de la forma $p_{i\alpha}^2/m_i$ y obtener

$$\left\langle \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{m_i} \right\rangle = 2 \langle \mathcal{K} \rangle = 3Nk_B T. \quad (6.68)$$

Este es el conocido principio de equipartición: una energía promedio de $k_B T/2$ por cada grado de libertad.

Es conveniente de definir una temperatura cinética instantánea

$$T = \frac{2\mathcal{K}}{3Nk_B} = \frac{1}{3Nk_B} \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{m_i} \quad (6.69)$$

cuyo promedio es igual a T . Esta definición no es única, por ejemplo para sistemas de moléculas rígidas, pueden aparecer velocidades angulares en la definición de T .

No debemos olvidar que para un sistema con N átomos, sujeto a N_r restricciones internas independientes (cómo ángulos internos fijos, cantidades conservadas, etc), el número de grados de libertad ya no será $3N$, sino $3N - N_r$, y debemos reemplazar la ecuación 6.69 por

$$T = \frac{2\mathcal{K}}{(3N - N_r)k_B} = \frac{1}{(3N - N_r)k_B} \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{m_i}. \quad (6.70)$$

Dado que usualmente hacemos cero el momento lineal del centro de masa, incluimos tres restricciones en la dinámica.

Si simulamos un ensamble en el que la temperatura sea una de las constantes podemos ver variaciones importantes en su valor. Estas variaciones en la temperatura pueden representar un problema para comparar los resultados con experimentos, que normalmente se realizan a temperatura y/o presión constantes. Para solucionar este problema podemos recurrir a alguno de los procedimientos descritos en la sección 6.14 para simular otros ensambles.

6.6.8 Curva calórica

Una vez que se han medido la energía total \mathcal{E} y la temperatura T para diferentes estados termodinámicos, se puede construir la *curva calórica* $\mathcal{E}(T)$. Esta es una herramienta útil para monitorear la aparición de transiciones de fase que implican un salto de $\mathcal{E}(T)$ si son de primer orden [Erc97], o de alguna derivada de $\mathcal{E}(T)$ si son de segundo o mayor orden. A pesar de esto, la curva calórica no siempre es una buena manera de estimar la temperatura de fusión T_m (la m por el inglés *melting*) cómo explico más adelante en la sección 6.7.

La transición de fase de primer orden más común para estudiar con dinámica molecular es la fusión, fácilmente observable en una simulación. Cuando el sistema abandona la estructura cristalina y se vuelve un líquido desordenado, observamos un salto de $\mathcal{E}(T)$, correspondiente a la energía de calor latente. Esto ocurre a una temperatura algo mayor que la verdadera temperatura de fusión T_m del modelo, debido a efectos de histéresis asociados con la necesidad de esperar a que aparezca una "semilla" de la fase líquida por las fluctuaciones espontáneas. Sólo cuando ha aparecido una "semilla" de algunos átomos puede la fase líquida comenzar a crecer a costa de la fase sólida. En la escala temporal típica de una simulación de dinámica molecular uno debe pasarse de la temperatura de fusión por 20 ó 30% para poder observar la fusión ocurriendo [Erc97, pág. 27].

En la sección 6.7 trato con mayor detalle lo referente a la medición de la temperatura de fusión.

6.6.9 Densidad

Calcular la densidad es relativamente simple y se puede hacer directamente de su definición:

$$\rho = \frac{N}{V} \quad (6.71)$$

en donde, V es el volumen del sistema y N el número de partículas contenidas en él. No siempre es trivial definir el volumen que estamos tratando (por ejemplo en cúmulos).

Si estamos tratando con un sistema periódico es más simple porque podemos referirnos al volumen de la celda original y a la cantidad total de partículas (átomos o moléculas) simuladas. También es posible hablar de la densidad específica para algún subconjunto, lo que resulta especialmente útil para observar fenómenos de difusión, fusión localizada y otros.

Cuando el sistema simulado es un líquido o gas no podemos estar interesados en colocar al sistema en un estado con una cierta densidad ρ . En este caso utilizamos la definición a la inversa y encontramos que

$$L = \left(\frac{N}{\rho} \right)^{\frac{1}{3}} \quad (6.72)$$

donde L es la longitud de un lado de la celda original que se necesita para obtener tal densidad. La ecuación 6.72 supone una celda original cúbica, en otros casos hay que adaptar la ecuación en forma especial.

Así como en las secciones 6.12 y 6.13 mostramos como controlar la temperatura y presión respectivamente, en una simulación estándar se puede controlar la densidad modificando el volumen de la celda de simulación, para poder mantener la cantidad de partículas constante. Normalmente no es necesario modificarlo drásticamente, y son suficientes unos puntos porcentuales [Erc97].

Fracción de empaque

Además de ρ podemos medir la fracción de empaque η , que es la razón entre el volumen de N esferas ($V_{esferas}$) y el volumen que las contiene (V)

$$\eta = \frac{V_{esferas}}{V} = \frac{N\sigma^3\pi}{6V} = \frac{\pi\rho}{6} \quad (6.73)$$

Medir la fracción de empaque es particularmente útil cuando utilizamos potenciales discontinuos, como el de esferas duras 7.4.2.

Existe un máximo para la fracción de empaque, que depende del potencial. No es fácil definirlo para potenciales suaves, pero para esferas duras sí. Para esferas duras el máximo ocurre cuando las esferas se arreglan ya sea en una estructura hexagonal (HCP) o en una cúbica (CCP), en donde tenemos

$$\eta_{\max} = \frac{\pi}{6}\sqrt{2} = 0.74048 \quad (6.74)$$

Según Allen [All87] "Muchos matemáticos creen, y todos los físicos saben, que la ecuación 6.74 da la máxima densidad posible de esferas duras uniformes", aunque aún no ha sido probada rigurosamente.

6.6.10 Presión

Un formalismo común para derivar una expresión para calcular la presión es el siguiente ([All87]):

Si elegimos coordenadas cartesianas, como es costumbre en dinámica molecular, y usamos las ecuaciones de movimiento de Hamilton, es fácil ver que la derivada respecto a cada coordenada es la negativa de una componente de la fuerza f_i ; sumando sobre las N moléculas

$$-\frac{1}{3} \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i^{ext} \right\rangle = -\mathcal{P}V. \quad (6.75)$$

Si definimos el *virial interno* \mathcal{W} como

$$-\frac{1}{3} \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_{\mathbf{r}_i} \mathcal{V} = \frac{1}{3} \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i = \mathcal{W} \quad (6.76)$$

y restringimos nuestra atención a las fuerzas intermoleculares obtenemos

$$\mathcal{P}V = Nk_B T + \langle \mathcal{W} \rangle. \quad (6.77)$$

Esto sugiere que, análogamente a como hicimos con la temperatura, definamos una función de presión instantánea [Che77]

$$\mathcal{P} = \rho k_B T + \frac{\mathcal{W}}{V} = \mathcal{P}^{id} + \mathcal{P}^{ex} \quad (6.78)$$

donde \mathcal{P}^{id} es una presión interna y \mathcal{P}^{ex} es más bien producto de un trabajo externo. El promedio de \mathcal{P} es simplemente P . Nuevamente esta definición no es única, pues más allá de las diferentes opciones al definir \mathcal{W} , puede ser más conveniente (por ejemplo en un ensamble a temperatura constante) usar

$$\mathcal{P}' = \rho k_B T + \frac{\mathcal{W}}{V} = \langle \mathcal{P}^{id} \rangle + \mathcal{P}^{ex} \quad (6.79)$$

Tanto \mathcal{P} como \mathcal{P}' promediadas dan P , pero sus fluctuaciones en general serán diferentes.

Erpenbeck y Wood mostraron en 1977 [Erp77] que la derivación anterior no es estrictamente válida para un sistema periódico infinito como los que intentamos simular en la computadora, pues no hay fronteras ni fuerzas externas.

Para interacciones de pares, \mathcal{W} puede expresarse en forma más simple, que sea explícitamente independiente del origen de coordenadas. Esto lo hacemos escribiendo \mathbf{f}_i como la suma de fuerzas \mathbf{f}_{ij} en el átomo i debido al átomo j

$$\sum_i \mathbf{r}_i \cdot \mathbf{f}_i = \sum_i \sum_{j \neq i} \mathbf{r}_i \cdot \mathbf{f}_{ij} = \frac{1}{2} \sum_i \sum_{j \neq i} \mathbf{r}_i \cdot \mathbf{f}_{ij} + \mathbf{r}_j \cdot \mathbf{f}_{ji}. \quad (6.80)$$

La segunda relación se debe a que los índices i y j son equivalentes. Usamos ahora la tercera ley de Newton $\mathbf{f}_{ji} = -\mathbf{f}_{ij}$ para intercambiar los índices de las fuerzas

$$\sum_i \mathbf{r}_i \cdot \mathbf{f}_i = \frac{1}{2} \sum_i \sum_{j \neq i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} = \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \quad (6.81)$$

en donde $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, y la última forma es normalmente la más conveniente. Es esencial usar la forma $\mathbf{r}_{ij} \cdot \mathbf{f}_{ij}$ en una simulación que use condiciones de frontera periódicas. Finalmente tenemos

$$\mathcal{W} = \frac{1}{3} \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} = -\frac{1}{3} \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot \nabla_{\mathbf{r}_{ij}} v(\mathbf{r}_{ij}) \quad (6.82)$$

$$= -\frac{1}{3} \sum_i \sum_{j > i} w(\mathbf{r}_{ij}) \quad (6.83)$$

en donde la función del virial de pares intermolecular $w(r)$ es

$$w(r) = r \frac{dv(r)}{dr}. \quad (6.84)$$

Cómo \mathcal{V} , \mathcal{W} está limitado por el alcance de las interacciones, y por lo tanto $\langle \mathcal{W} \rangle$ debe ser bien comportado e independiente del ensamble en la mayoría de los casos ([All87]).

No es la única, y me parece correcto incluir en el apéndice D.1 otro formalismo.

Tal y cómo hicimos para la energía potencial al agregar una función de corrección de cola para las interacciones más allá del radio de corte, podemos hacerlo para la presión. Heeman [Hee90] hace una derivación de la ecuación de corrección de cola

$$\mathcal{P} = \frac{\rho^2}{6} \int_{r_c}^{\infty} g(r) \frac{\partial \phi}{\partial r} 4\pi r^3 dr. \quad (6.85)$$

y explica que pueden llegar a significar varios puntos porcentuales del valor total.

El cálculo de la presión en sistemas bajo acción de potenciales como el de esferas duras 7.4.2 es diferente, y sugiero revisar el texto de Allen [All87, sección 5.6] para un tratamiento.

6.6.11 Desplazamiento cuadrático medio

El desplazamiento cuadrático medio de los átomos en una simulación puede ser calculado fácilmente según su definición

$$DCM \equiv \langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle \quad (6.86)$$

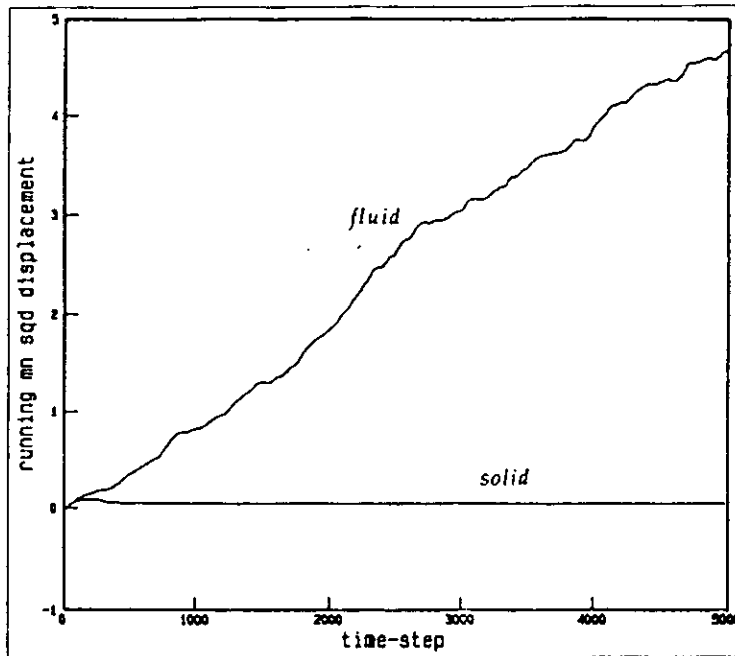


Figura 6-15 Gráfica del desplazamiento cuadrático medio móvil, que permite distinguir un sólidos de líquidos. En la simulación se usaron 256 átomos, con la misma densidad $\rho = 0.9$ pero con temperaturas $T = 0.8$ y $T = 1.087$ respectivamente. [Hai92]

en donde $\langle \cdot \rangle$ denota el promedio sobre los átomos (todos o algún subconjunto). Hay que tener cuidado para evitar considerar “saltos” de partículas al utilizar las condiciones de frontera periódicas.

La DCM es útil porque contiene información de la difusividad atómica.

Experimentalmente se observan diferencias importantes en el comportamiento de la DCM según el estado del sistema observado:

- En un sólido la DCM debe ser muy cercano a cero, porque los movimientos principales de los átomos son de oscilación alrededor de sus posiciones de equilibrio.
- En un líquido la DCM aumenta linealmente con el tiempo por la deriva de los átomos alejándose de su posición original.
- En un gas la DCM aumenta en forma cuadrática con el tiempo (ver la sección 6.10.1, de las funciones de autocorrelación espacio-tiempo, doy una explicación de porque proporcional a t^2).

No he hecho ningún esfuerzo aquí por explicar las diferencias de comportamiento de la DCM para diferentes estados. En la figura 6-15 se muestra el promedio móvil de la DCM contra el tiempo para un sistema de Lennard-Jones.

Si la pendiente de la DCM depende del estado del sistema, podemos definir el *coeficiente de difusión* en función de ella:

$$D \equiv \lim_{t \rightarrow \infty} \frac{1}{6t} \langle |r(t) - r(0)|^2 \rangle \quad (6.87)$$

El 6 en la fórmula anterior corresponde con dos veces el número de dimensiones, y debe sustituirse por un 4 si estamos tratando con un sistema bidimensional o un 2 en sistemas unidimensionales (sección 5.3.1).

Rapaport [Rap95] presenta la ecuación de Green-Kubo para la difusión basada en la función de autocorrelación de velocidades:

$$D = \frac{1}{3N_a} \int_0^\infty \langle \sum_{j=1}^\infty v_j(t) \cdot v_j(0) \rangle dt, \quad (6.88)$$

que según él es completamente equivalente a la ec. 6.87.

La figura 6-14 muestra las diferencias en trayectoria para sistemas en estado sólido y líquido, haciendo evidente la diferencia en su DCM. Es común modificar ligeramente la forma de calcular este coeficiente utilizando los llamados “promedios móviles”. Estos consisten en ir sustituyendo el estado considerado como inicial conforme avanza la

simulación, de esta forma se elimina el carácter de *especial* que se le otorga a la condición inicial, y se obtienen curvas más suaves que las de la figura 6-14. Su implementación es ligeramente más complicada y requieren más espacio de memoria, pero si buscamos el coeficiente de difusión como derivada del DCM vale la pena.

6.6.12 Energía libre

A pesar de poder agregar errores y problemas, el tamaño finito del sistema tiene ventajas, por ejemplo, permite calcular propiedades termodinámicas de segundo orden. En un sistema de tamaño finito las propiedades intensivas que describen el sistema se desvían de sus valores medios, fluctuando alrededor de estos. Podemos utilizar, por ejemplo, la definición de calor específico C_V ,

$$C_V = \frac{\partial}{\partial T} \left[T^2 \frac{\partial \left(\frac{F}{T} \right)}{\partial T} \right] \quad (6.89)$$

y su relación con las fluctuaciones en la temperatura del ensamble microcanónico [Hee90] (sección 6.14)

$$\frac{\langle T^2 \rangle - \langle T \rangle^2}{\langle T \rangle^2} = \frac{3}{2N} \left[1 - \frac{3kN}{2C_V} \right] \quad (6.90)$$

para calcular la energía libre F , remediando la dificultad para calcular la energía libre directamente.

6.7 Midiendo la temperatura de fusión

El comportamiento de la desviación cuadrática media (DCM) como función del tiempo puede permitirnos distinguir entre sólidos y líquidos. Uno puede entonces ser tentado a localizar la temperatura de fusión T_m de la sustancia simulada incrementando la temperatura de un sistema cristalino hasta observar difusión o un salto en la curva calorica mostrando la absorción de calor.

A pesar de que estos son indicadores de una transición de sólido a líquido, la temperatura a la que ocurren es invariablemente mayor que la temperatura de fusión correcta. De hecho, el punto de fusión es por definición la temperatura a la que la fase sólida y líquida coexisten (tienen la misma energía). A falta de una “semilla” desde la que pueda extenderse la fase líquida, sobrecalentamos por encima del punto de fusión. En esta región el sistema se encuentra en estados termodinámicos metaestables, sin embargo es aparentemente estable en la escala temporal de la simulación. Un cristal sobrecalentado se rompe cuando se alcanza su punto de estabilidad mecánica. Este punto puede ser mayor a T_m por un 20-30%.

Mientras la inestabilidad mecánica es ciertamente útil en estimar aproximadamente la localización de T_m , son necesarios, y existen, métodos más precisos. Uno de ellos consiste en colocar una muestra grande, consistente de aproximadamente 50% sólido y 50% líquido. Una muestra como esta contendrá regiones de interface entre sólido y líquido. Ahora podemos tratar de establecer un estado de equilibrio en el que coexistan sólido y líquido. Cuando consigamos esto, hemos encontrado, por definición, la temperatura de fusión T_m . Una manera de construir un sistema como el propuesto es no permitir el movimiento (“congelar”) de la mitad de las partículas, mientras elevamos la temperatura fundiendo la otra mitad, para finalmente liberar las primeras y poder observar el desarrollo. Hay que tener cuidado de no sobrecalentar el grupo que fundimos, pero si conocemos aproximadamente la temperatura de fusión no tendremos problemas en conseguir esto.

Un sistema generado según la receta anterior tendrá una temperatura inicial T_0 que esperamos sea cercana a la del punto de fusión T_m . Supongamos que $T_0 > T_m$, entonces, la fase líquida es favorecida por encima de la sólida, y las interfaces sólido-líquido existentes comienzan a moverse incrementando la fracción del material que es líquido a costa del sólido. En otras palabras, algo del material se funde. Mientras esto ocurre el calor latente de fusión es absorbido por el sistema. Como la energía total del sistema se conserva, la absorción de calor latente automáticamente implica una disminución de la energía cinética. Por lo tanto, la temperatura baja. La “fuerza” que controla el movimiento de la interface sólido-líquido también debe disminuir, y conforme transcurra el tiempo, la posición de la interfase y la temperatura han de converger exponencialmente hacia un estado de equilibrio. La temperatura será entonces T_m .

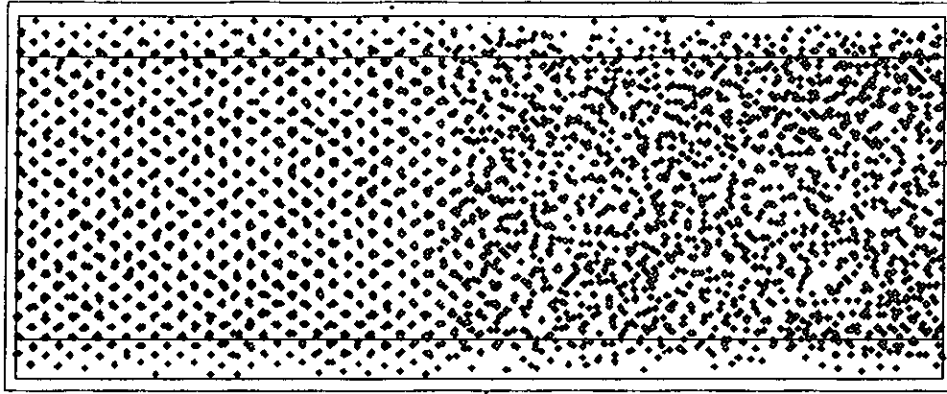


Figura 6-16 Para medir la temperatura de fusión podemos construir un sistema con fases sólida (izquierda) y líquida (derecha). Los átomos han sido marcados gris oscuro si su desplazamiento respecto a la posición inicial es poco (menor a cierto valor) y con gris claro si es mucho (mayor). Modificando suavemente las condiciones de temperatura y presión podemos mover la línea de frontera entre las fases y encontrar con precisión la temperatura de fusión. [Ame99]. Comparese con la figura 6-14 en donde se muestran las trayectorias en sistemas aislados líquido y sólido.

Si por el contrario $T_0 < T_m$, el razonamiento inverso es aplicable: es liberado calor latente y convertido en energía cinética. La temperatura nuevamente converge exponencialmente a T_m , esta vez desde abajo.

La figura 6-16 [Ame99] es un ejemplo de una simulación de este tipo. A la izquierda se distingue una parte del sistema en estado sólido y estructura cristalina, a la derecha estado líquido con átomos desordenados. Coloreando los átomos con oscuro y claro dependiendo de cuanto se han desplazado de cierta posición inicial es fácil distinguir ambas fases. Es interesante notar la zona intermedia, en donde se observan átomos de ambos colores en una configuración aproximadamente cristalina. Las delgadas líneas horizontales representan las regiones en que se pueden hacer mediciones confiables sin preocuparnos por efectos de borde. Comparese con la figura 6-14 en donde se muestran las trayectorias en sistemas aislados líquido y sólido.

La temperatura de fusión depende de la presión del sistema. Una medida de la presión en el estado final es necesaria para caracterizarlo termodinámicamente. Es normal utilizar el procedimiento sugerido utilizando técnicas de simulación a presión constante (sección 6.14.3).

6.8 Correlaciones en el espacio real

Las funciones de correlación en el espacio real tienen típicamente la forma

$$\langle A(\mathbf{r})A(\mathbf{0}) \rangle \quad (6.91)$$

y son directamente obtenibles de la dinámica molecular.

El ejemplo más simple es la función de correlación de pares $g(r)$, que es la probabilidad de encontrar un par de átomos separados por una distancia r , relativo a lo que esperamos en una distribución uniforme con la misma densidad de partículas. Matemáticamente:

$$g(r) \equiv \frac{1}{\rho^2 N} \left\langle \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \delta(r - r_{ij}) \right\rangle \quad (6.92)$$

Esta función lleva información de la estructura del sistema. Para un cristal, exhibe una serie de picos en las posiciones correspondientes a las capas de vecinos alrededor de un átomo. Las posiciones y tamaños de los picos son una huella de la estructura del cristal.

Para los líquidos $g(r)$ exhibe su pico principal cerca del promedio de la separación atómica entre átomos vecinos, y posteriormente presenta oscilaciones menos pronunciadas. El tamaño de los picos disminuye exponencialmente con la distancia conforme $g(r)$ se aproxima a 1.

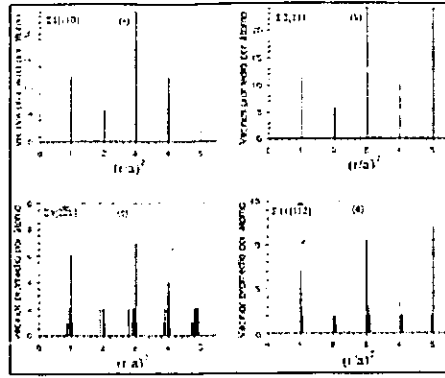


Figura 6-17 Función de distribución radial para diferentes estructuras. a) Cristal FCC perfecto. Fronteras de inclinación según el modelo CSL con b) $\Sigma = 3$, c) $\Sigma = 9$ y d) $\Sigma = 11$.

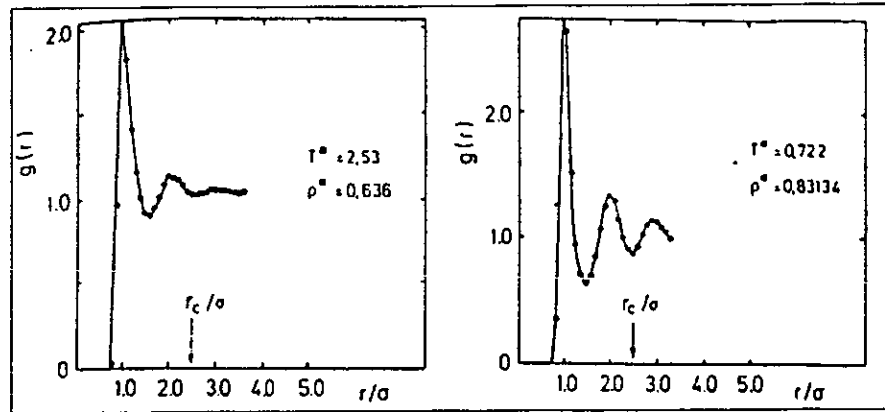


Figura 6-18 Funciones de correlación de pares para dos estados diferentes. [Hee90]

En sólidos y líquidos $g(r)$ es 0 para distancias pequeñas, para las que la repulsión atómica es tan grande como para evitar que los átomos se aproximen demasiado.

Una cantidad comúnmente calculada con $g(r)$ es la *función de distribución radial*, que es el número promedio de átomos localizados entre las distancias r_1 y r_2 de un átomo cualquiera,

$$\rho \int_{r_1}^{r_2} g(r) 4\pi r^2 dr \quad (6.93)$$

Esta cantidad permite definir números de coordinación incluso en sistemas desordenados.

El cálculo de $g(r)$ requiere naturalmente un número de cálculos de orden $O(N^2)$ y por lo tanto puede retrasar considerablemente un programa optimizado. Si el comportamiento a grandes r no es importante, puede ser conveniente definir una distancia de corte. Al hacer los cálculos de las interacciones se tienen disponibles las distancias interatómicas, por lo que se acostumbra aprovecharlas para generar la función de distribución de pares.

La figura 6-17 muestra la función de distribución radial (número de vecinos promedio por átomo como función de la distancia) para un cristal FCC perfecto y diferentes configuraciones de fronteras CSL (sección 1.3.5).

Se puede hacer notar también que usar condiciones de frontera periódicas impone una distancia de corte máxima $L_m/2$, en donde $L_m = \min\{L_x, L_y, L_z\}$ en todas las direcciones. A distancias mayores los resultados se contaminan por la periodicidad impuesta al sistema y dejan de tener relevancia.

La figura 6-18 muestra las funciones de correlación de pares en puntos distintos del diagrama de fase del argón de Lennard-Jones.

Siendo $g(r)$ independiente del tiempo también es común hacer promedios temporales.

6.9 Correlaciones del espacio recíproco

No basta con observar la evolución de la energía y el momento para saber la forma en que evoluciona un sistema simulado, necesitamos también cantidades que puedan medir propiedades estructurales. Para un sistema con densidad espacial

$$\rho(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \quad (6.94)$$

la más básica de las cantidades es la *transformada de fourier*

$$\rho(\mathbf{k}) = \sum_{i=1}^N e^{i\mathbf{k} \cdot \mathbf{r}_i} \quad (6.95)$$

que se obtiene rápida y fácilmente de una configuración, pero que constituye la clave para construir el *factor de estructura estático*

$$S(\mathbf{k}) = \frac{1}{N} \langle \rho(\mathbf{k}) \rho(-\mathbf{k}) \rangle \quad (6.96)$$

$$= \frac{1}{N} \left\langle \sum_{i,j}^N e^{i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)} \right\rangle \quad (6.97)$$

$$= \frac{1}{N} \left\langle \sum_{i,j}^N \cos(\mathbf{k} \cdot \mathbf{r}_{ij}) \right\rangle \quad (6.98)$$

que permite comparar resultados con experimentos de dispersión [Erc97]. En líquidos, que son isotrópicos, $S(k)$ depende únicamente del módulo del vector de onda, y es en esencia una transformada de fourier de la función de correlación de pares

$$S(k) = 1 + \rho g(k). \quad (6.99)$$

En ocasiones el cálculo de $S(k)$ con 6.96 es mucho más rápido, y preferible, que el cálculo de $g(k)$.

Remplazando $\rho(\mathbf{k})$ por alguna otra observable $A(\mathbf{k})$ podemos igualmente calcular cualquier otra función de correlación en el espacio recíproco.

Si estamos trabajando con un sistema periódico nos basta con conocer la localización de las partículas en la celda original para conocer el resto. De esta forma podemos imponer las condiciones (por ejemplo sobre la dirección x)

$$e^{i\mathbf{k}_x(x_1 + L_x)} = e^{i\mathbf{k}_x x_1} \quad (6.100)$$

en donde L_x es el tamaño de la celda de reptición. Esto implica una restricción en las posibilidades al elegir el vector de onda, quedando restringidos a

$$k_x = \frac{2\pi}{L_x} n_x \quad (6.101)$$

con n_x entero. Igualmente ocurre para y y z , dando lugar a una *cuantización* de los vectores de onda posibles. En estos casos $\rho(k)$ recibe el nombre de *parámetro de orden* y se acostumbra denotar por λ .

Dadas sus raíces en la periodicidad del sistema, el parámetro de orden es particularmente útil al tratar con estructuras cristalinas y de muy poca utilidad si estudiamos sistemas gaseosos. Para cristales FCC se reduce a (incluyendo ahora las tres coordenadas espaciales)

$$\lambda = \frac{1}{3} [\lambda_x + \lambda_y + \lambda_z] \quad (6.102)$$

en donde

$$\lambda_x \equiv \frac{1}{N} \sum_i^N \cos\left(\frac{4\pi x_i}{L_x}\right). \quad (6.103)$$

Si inicialmente los átomos de un sistema se localizan en los puntos de una red FCC, $\lambda_{inicial} = 1$. porque todos los componentes de las posiciones atómicas x_i , y_i y z_i son múltiplos de $\frac{L}{2}$. Cuando la red se ha desintegrado por completo λ ha de fluctuar alrededor de cero, porque los átomos están uniformemente distribuidos alrededor de sus puntos de red. Las fluctuaciones serán de magnitud $\frac{\sqrt{N}}{N}$.

Es posible definir para otras estructuras regulares un parámetro de orden análogo a 6.103.

En la sección 6.10.3 describo el factor de estructura dinámico, más general(y por supuesto más complicado).

6.10 Análisis dinámico

“Lo más difícil de entender en el mundo es el iva.” Albert Einstein.

Una de las virtudes de la dinámica molecular con respecto a los métodos de simulación estocásticos (como Monte Carlo) es poder seguir la evolución temporal del sistema, y por lo tanto, tener información acerca de la dinámica del sistema. Datos experimentales, como curvas de dispersión de fonones en el caso de cristales, se pueden obtener del espacio de frecuencias del vector de onda [Erc97]. Información obtenida por dinámica molecular en el espacio y tiempo reales pueden ser transformadas mediante Fourier y comparadas así con resultados experimentales.

Una cantidad típica a calcular es la función de correlación de Van Hove $G(\mathbf{r}, t)$, para entenderla me gustaría citar el texto de Haile [Hai92, pág. 310]:

Imagine llegar al final de las escaleras que dan al gran salón de una estación de trenes, es viernes por la tarde, hora pico. Desde su posición en la escalera hay que arreglarselas, a través de la multitud, hasta la taquilla, que se encuentra a una distancia r_a de las escaleras, y que estimas te tomara un tiempo t_a recorrer. El tren se encuentra a una distancia r_b de las escaleras, Estimas tomara otros t_b minutos comprar el boleto y caminar al andén. Debido a la multitud tu avance a través de la estación será no lineal y te preguntas a) ¿cual es la probabilidad $P_d(r_a, t_a)$ de que haya otra persona en la taquilla cuando llegues? y b) ¿cual es la probabilidad $P_s(r_b, t_a + t_b)$ de que llegues al andén en $t_a + t_b < 5$ minutos.

Las cantidades P_s y P_d reflejan información acerca de la estructura de la multitud; en particular, estas propiedades dependen de cómo la densidad local $\rho(\mathbf{r}, t)$ se relaciona con la densidad local $\rho(0, 0)$. Para asuntos a nivel molecular, estas probabilidades estructurales están relacionadas con una función de correlación temporal densidad-densidad $G(\mathbf{r}, t)$, que está definida por

$$G(\mathbf{r}, t) \equiv \frac{1}{\rho} \langle \rho(\mathbf{r}, t) \rho(0, 0) \rangle = \frac{1}{N} \left\langle \sum_i \sum_j \delta[\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}_j(t)] \right\rangle \quad (6.104)$$

en donde δ es la Delta de Dirac y los paréntesis $\langle \cdot \rangle$ significan, como siempre, el promedio sobre el ensamble. La cantidad $G(\mathbf{r}, t)$ se llama *función de correlación espacio-tiempo*. Es el análogo clásico de la función de van Hove usada para caracterizar la estructura dinámica medida en experimentos de dispersión de neutrones. Notese de 6.104 que G tiene unidades de densidad, y por lo tanto la δ en 6.104 debe tener unidades de V^{-1} . Consecuentemente $G(\mathbf{r}, t)$ no es por sí misma una probabilidad, sino que es proporcional a una probabilidad.

$G(\mathbf{r}, t) \propto$ probabilidad de que un átomo esté en la posición \mathbf{r} al tiempo t dado que un átomo estaba en el origen $\mathbf{r} = 0$ al tiempo inicial $t = 0$.

La función de correlación espacio-tiempo es una cantidad dinámica. Sin embargo es una propiedad de equilibrio porque se obtiene al promediar sobre una trayectoria de equilibrio en el espacio fase.

Es evidente de 6.104 que la función de correlación espacio-tiempo puede separarse en dos partes. Términos con $i = j$ dan lugar a la *función de auto-correlación* $G_s(\mathbf{r}, t)$ (s por el inglés de *self*), para la que un átomo en (\mathbf{r}, t) es el mismo átomo que ocupaba el origen ($\mathbf{r} = 0, t = 0$). Términos con $i \neq j$ dan lugar a la función de correlación distintiva $G_d(\mathbf{r}, t)$ para la que el átomo en (\mathbf{r}, t) es diferente del que ocupaba el origen. Podemos entonces escribir 6.104 como

$$G(\mathbf{r}, t) = G_s(\mathbf{r}, t) + G_d(\mathbf{r}, t). \quad (6.105)$$

La función de auto-correlación la trato en la siguiente subsección y la función de correlación distintiva queda para la subsección 6.10.2.

6.10.1 Función de auto-correlación espacio-tiempo

Si utilizamos la definición de Haile de la función de autocorrelación de la ecuación 6.104 encontramos

$$G_s(\mathbf{r}, t) = \frac{1}{N} \left\langle \sum_i \delta[\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}_i(t)] \right\rangle \quad (6.106)$$

Si G_s depende solamente de r , como en sistemas homogéneos, y no de la dirección, podemos escribir

$$G_s(r, t) = \frac{1}{N} \left\langle \sum_i \delta[r - |\mathbf{r}_i(0) - \mathbf{r}_i(t)|] \right\rangle. \quad (6.107)$$

Podemos resumirla así:

$$G_s(r, t)dr = \begin{cases} \text{probabilidad de que al tiempo } t \text{ un átomo este a una distancia } r \\ \text{del origen } (r = 0), \text{ dado que el mismo átomo} \\ \text{estaba en el origen al tiempo inicial } t = 0. \end{cases} \quad (6.108)$$

Si $G_s(r, t)$ representa una probabilidad debemos pedir que se encuentre normalizada para cada instante

$$\int_0^t G_s(r, t)dr = 1 \quad (6.109)$$

con la condición de que si $t = 0$ obtengamos

$$G_s(r, 0) = \delta. \quad (6.110)$$

Para átomos cuyo movimiento no este relacionado con su posición inicial podemos decir: para tiempos largos es igualmente probable encontrarlos en un volumen V (no necesariamente fácil de determinar) y G_s será una constante. Para encontrar el valor de la constante bastara con normalizar usando la ecuación 6.109. Si el volumen en que se puede encontrar al átomo aumenta con el tiempo (como es de esperarse y ocurre en la mayoría de los casos), encontramos

$$\lim_{t \rightarrow \infty} G_s(r, t) = \lim_{r \rightarrow \infty} G_s(r, t) \times \frac{1}{V} \approx 0 \quad (6.111)$$

El comportamiento de $G_s(r, t)$ sera tal que comenzando cómo una delta colapse a casi cero al avanzar el tiempo. Según Haile durante el colapso podemos esperar que se comporte como una gaussiana. Como G_s describe el movimiento de un solo átomo, podemos esperar que su movimiento este relacionado con la desviación cuadratica media de ese átomo, y poder usar la aproximación

$$G_s(r, t) \approx \frac{1}{\left[\frac{4}{3}\pi \langle \Delta r^2(t) \rangle\right]^{\frac{3}{2}}} e^{-\frac{r^2}{\frac{4}{3}\pi \langle \Delta r^2(t) \rangle}} \quad (6.112)$$

en donde $\langle \Delta r^2(t) \rangle$ es la desviación cuadratica media al tiempo t . Para fluidos la gaussiana 6.112 es una aproximación confiable para distancias pequeñas y tiempos largos. Si los tiempos son pequeños (antes de que ocurran colisiones), un átomo debe moverse como una partícula libre, y las trayectorias seran líneas rectas $r = vt$, y $\langle \Delta r^2(t) \rangle$ estara dada por

$$\langle \Delta r^2(t) \rangle = \frac{3kT}{m} t^2 \quad (6.113)$$

Para tiempos grandes, después de muchas colisiones, el movimiento atómico es más bien parecido a una caminata al azar, obedeciendo la relación de Einstein

$$\lim_{t \rightarrow \infty} \frac{\langle [r(t) - r(0)]^2 \rangle}{6t} = D \quad (6.114)$$

en donde D es el coeficiente de auto-difusión.

$$\langle \Delta r^2(t) \rangle = 6D \quad (6.115)$$

La aproximación gaussiana no sera muy confiable a tiempos intermedios. La duración del período intermedio depende en las condiciones de estado [Rah64] [Ste80]. Un ejemplo de G_s es la figura 6-19 en la que los puntos son resultados de simulaciones y las líneas cálculos de la gaussiana de la ecuación 6.112 usando valores para la desviación cuadrática media tomados de la simulación. Para la condición de estado de la figura 6-19, la aproximación gaussiana no es confiable para tiempos entre 0.5 y 2 ps.

Haile [Hai92] obtiene la relación

$$G_s(r, t) = \frac{1}{MNV(r, \Delta r)} \sum_k \sum_i \delta[r - |r_i(t_k) - r_i(t_k + t)| \Delta r]_{\Delta r} \quad (6.116)$$

en donde el índice k corre sobre el total M de orígenes de tiempo.

Por lo tanto, calcular $G_s(r, t)$ involucra propiedades de la función de distribución radial combinadas con propiedades de las funciones de correlacion temporales. Usualmente un valor de $\Delta r = 0.025\sigma$ es apropiado [Hai92, sección 6.4.1].

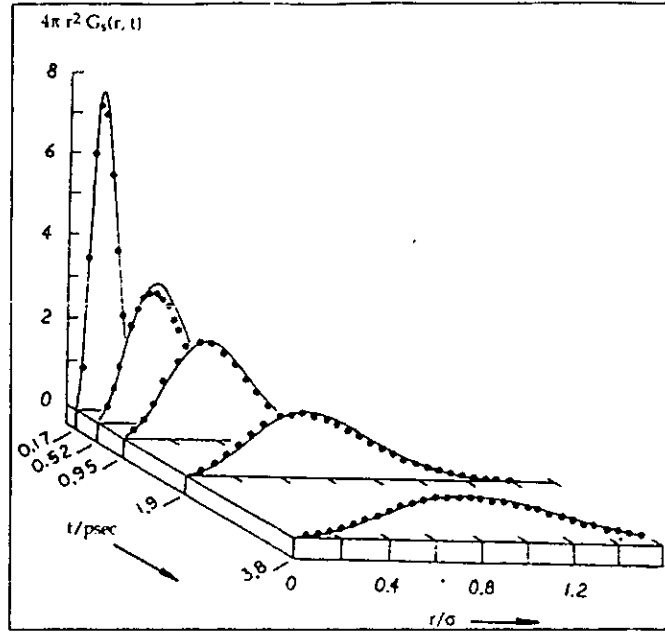


Figura 6-19 Función de autocorrelación de espacio tiempo para un sistema de 256 átomos interactuando según un potencial de Lennard-Jones. Los puntos son resultados de simulaciones obtenidos de promedios sobre 2000 orígenes de tiempo, cada uno separado 10dt . Las líneas son la aproximación gaussiana usando valores del desplazamiento cuadrático medio.[Hai92]

6.10.2 Función de correlación distintiva espacio-tiempo

Según la ecuación 6.104, para un sistema homogéneo uniforme la función de correlación distintiva espacio-tiempo es

$$G_d(r,t) = \frac{1}{N} \left\langle \sum_i \sum_j \delta[|r - |r_i(0) - r_j(t)||] \right\rangle \quad (6.117)$$

y se puede expresar:

$$\frac{1}{N} G_d(r,t) dr = \begin{cases} \text{probabilidad de que al tiempo } t \text{ un átomo este a una distancia } r \\ \text{del origen (} r = 0 \text{), dado que algún otro átomo} \\ \text{estaba en el origen al tiempo inicial } t = 0. \end{cases} \quad (6.118)$$

La normalización es

$$\frac{1}{N} \int G_d(r,t) dr = 1 - \frac{1}{N} \quad (6.119)$$

En otras palabras, 6.119 quiere decir que acumular G_d sobre el volumen accesible simplemente cuenta los $N - 1$ átomos diferentes al átomo i del contenedor. Al tiempo inicial, la función estática $G_d(r,0)$ es proporcional a la función de distribución radial

$$G_d(r,0) = \rho g(r). \quad (6.120)$$

mientras que para tiempos y separaciones grandes la posición de un átomo no esta relacionada con la posición anterior de otro átomo, por lo que

$$\lim_{r \rightarrow \infty} G_d(r,t) = \lim_{t \rightarrow \infty} G_d(r,t) = \rho \quad (6.121)$$

Las ecuaciones 6.120 y 6.121 significan que conforme avanza el tiempo esperamos que el comportamiento de G_d cambie desde el valor inicial $\rho g(r)$ al valor final límite ρ para cualquier separación de pares r . La figura 6-20 muestra la evolución temporal de la función de correlación distintiva para un fluido de Lennard-Jones. En la figura 6-20 se observa que en 1 ps el primer máximo disminuye en 50% (de 2.5 a alrededor de 1.25); es decir, para este estado en particular los primeros vecinos no se mantienen alrededor de su sitio de inicio por mucho tiempo.

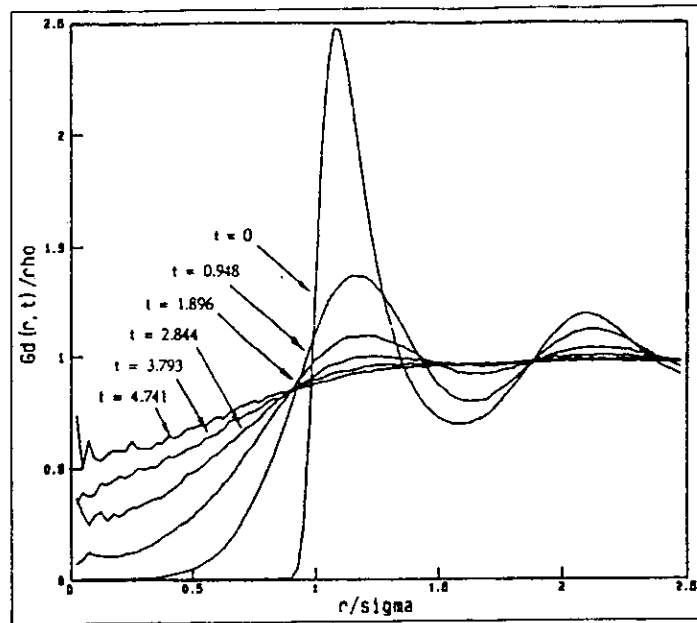


Figura 6-20 Función de correlación espacio-tiempo distintiva. Los promedios son sobre 2000 orígenes separados por $10dt$. Los tiempos t mostrados son en picosegundos. [Hai92]

Para obtener G_d en una simulación se puede derivar la expresión [Hai92]

$$G_d(r, t) = \frac{2}{MN\Delta r} \sum_k \sum_i \sum_{j>i} \delta[r - |r_i(t_k) - r_j(t_k + t)|] \Delta r \quad (6.122)$$

Un punto a considerar de la ecuación 6.122 es que debido a la triple suma sobre los átomos, y comparada con la doble suma de la ecuación 6.116 el cálculo de G_d requiere de aproximadamente un factor de N más tiempo de cálculo que el cálculo de G_s . Esta diferencia tiene como consecuencia que la función de correlación distintiva no sea calculada más que para sistemas de pocos átomos. Por ejemplo, para un sistema de 256 átomos para el que calcular G_s con 6.116 requiera de 1 minuto, el cálculo de G_d con 6.122 requerirá de 4 horas.

6.10.3 Factor de estructura dinámico

Así como la densidad espacial está relacionada con el factor de estructura estático (sección 6.9) la función de correlación de Van Hove está relacionada mediante una transformación de Fourier con el *factor de estructura dinámico* $S(\mathbf{k}, \omega)$ que reportan los experimentos [Erc97]:

$$S(\mathbf{k}, \omega) = \int \int e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} G(\mathbf{r}, t) d\mathbf{r} dt \quad (6.123)$$

Al igual que para cantidades estáticas, la forma más conveniente (en términos de tiempo de cómputo) es aquella que usa la transformada de Fourier de la densidad

$$\rho(\mathbf{k}, t) = \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i(t)} \quad (6.124)$$

Esta cantidad corresponde a una sola partícula y es rápida de calcular por contener sólo una suma sobre N . De ella uno construye la función de correlación densidad-densidad dependiente del tiempo (también llamada *función de dispersión intermedia*) haciendo un promedio sobre una trayectoria de dinámica molecular

$$F(\mathbf{k}, \omega) = \frac{1}{N} \langle \rho(\mathbf{k}, t + t_0) \rho(-\mathbf{k}, t_0) \rangle \quad (6.125)$$

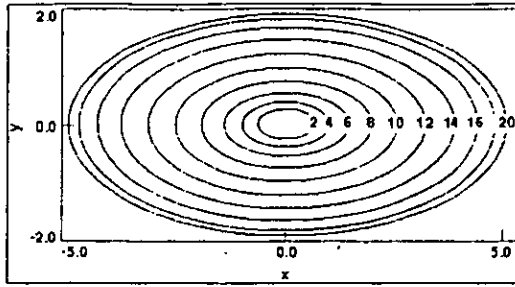


Figura 6-21 Curvas de nivel de la función de energía simple $\mathcal{E}(x, y) = x^2 + 5y^2$. Cada curva representa un aumento de la energía de dos unidades arbitrarias. [Msi97]

El promedio se hace sobre todas las posibles opciones para t_0 . El máximo t (que determina la resolución de frecuencias) esta limitado por la longitud de la simulación, mientras que la periodicidad con la que los datos se recolectan (que depende tanto del tamaño de δt cómo del número de intervalos entre cálculos) determina la máxima frecuencia en el espectro final. El factor de estructura dinámico se obtiene finalmente por una transformada de fourier temporal

$$S(\mathbf{k}, \omega) = \int F(\mathbf{k}, t) e^{i\omega t} d\mathbf{r} dt \quad (6.126)$$

Nuevamente los vectores \mathbf{k} permitidos se encuentran cuantizados cómo consecuencia de una periodicidad del sistema. Haciendo cálculos a diferentes temperaturas se puede estudiar, por ejemplo, el efecto de la temperatura en el espectro de fonones, incluyendo los efectos anarmónicos completos.

Al tratar con modos vibracionales y polarizaciones diferentes, es conveniente seguir lo anterior, pero utilizando corrientes

$$j(\mathbf{k}, t) = \sum_i \mathbf{v}_i(t) e^{i\mathbf{k} \cdot \mathbf{r}_i(t)} \quad (6.127)$$

en lugar de la densidad $\rho(\mathbf{k}, t)$.

6.11 Dinámica molecular como herramienta de optimización

Encontrar una configuración de equilibrio de un sistema significa encontrar una configuración tal que la fuerza neta sobre cada uno de los átomos sea cero. Basta con minimizar la energía para encontrar configuraciones que además de ser de equilibrio sean estables. Agregando fuerzas externas al sistema en forma de restricciones se puede responder a otras preguntas como ¿cuanta energía se requiere para que un sistema pase de una configuración a otra?, que puede responderse obligando a algunos átomos a encimarse con átomos de un templete durante una minimización de energía.

La eficiencia de un proceso de minimización se mide tanto por el tiempo empleado para encontrar un mínimo, cómo por la calidad del mínimo en el sentido de que tan amplio es el dominio para el que es mínimo.

Una manera de mostrar los diferentes mecanismos de minimización es aplicar cada algoritmo y capturar la esencia del proceso usando una función simple, como por ejemplo:

$$\mathcal{E}(x, y) = x^2 + 5y^2 \quad (6.128)$$

que describe una superficie elíptica de dos dimensiones.

Cualquier proceso de minimización comienza con una ecuación análoga a 6.128, que define una superficie de energías y un conjunto de condiciones iniciales (en este caso de coordenadas pero que en sistemas atómicos incluye las velocidades). La figura 6-21 muestra las curvas de nivel de la función 6.128 en el plano (x, y) , cada elipse esta separada dos unidades de energía y representa un conjunto de puntos con idéntica energía.

El mínimo de esta función localizado en $(0,0)$ se puede encontrar fácilmente por inspección, pero nuestra intención es hallarlo analíticamente.

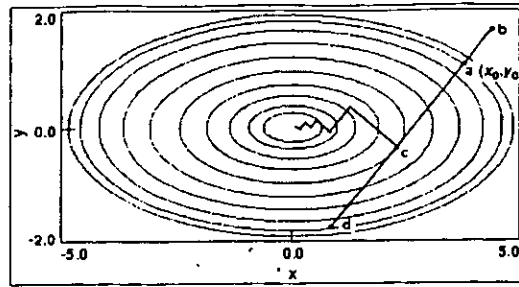


Figura 6-22 Curvas de nivel de energía de la ecuación 6.128. El vector derivada en el punto inicial $a(x_0, y_0)$ define la dirección para la búsqueda en línea. Notese que la derivada no apunta en la dirección del mínimo, y que el mínimo (punto c) ocurre justo en el momento en que la línea es tangente a la curva de nivel, lo que significa que las direcciones subsecuentes serán ortogonales. [Msi97]

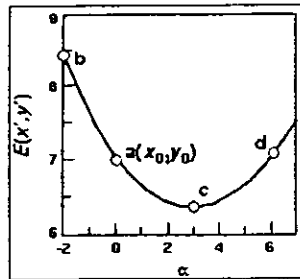


Figura 6-23 Corte de la superficie de energía de la ecuación 6.128, según la línea de búsqueda de la figura 6-22. [Msi97]

Dadas la función de energía y el punto de inicio, un minimizador debe determinar tanto la dirección hacia el mínimo como la distancia a este en esa dirección. Una buena dirección inicial es simplemente la pendiente (las derivadas) de la función de energía en el punto inicial. Las derivadas de la ecuación 6.128 son el vector bidimensional

$$\nabla \mathcal{E} = (2x, 10y) \quad (6.129)$$

En este caso las derivadas son proporcionales a las coordenadas, de tal manera que conforme uno se aleja del mínimo las derivadas crecen (ver figura 6-22). Sin embargo, simplemente apuntan hacia regiones de menor energía, y conforme nos movemos en esa dirección las derivadas cambian y apuntan en otra dirección. Para mejorar la eficiencia del método hay algoritmos un poco más sofisticados, como el gradiente conjugado y Newton-Raphson que aprovechan la información respecto a la forma en que cambian estas derivadas.

6.11.1 Búsqueda en línea

Antes de mostrar los diferentes algoritmos es útil introducir el concepto de *búsqueda en línea*, implícito en muchos métodos de minimización, que consiste en cambiar las coordenadas a unas correspondientes a una menor energía. Por ejemplo, en la figura 6-22 la dirección del gradiente de un punto de inicio arbitrario $a = (x_0, y_0)$ está superpuesto en la función.

La búsqueda en línea consiste en utilizar la derivada de la función de energía en un punto y construir una línea recta que vamos a parametrizar para construir una figura como la 6-23. Encontrando el mínimo de esta función unidimensional elegimos ese punto y utilizamos la recta ortogonal como nueva línea a parametrizar (ver figura 6-22). Utilizamos la recta ortogonal porque el mínimo en la curva de parametrizada coincide con el punto en que la curva de nivel de energía es paralela a la recta usada, lo que significa que la dirección de máximo cambio (su gradiente) es ortogonal.

Para el caso de la figura 6-22 el vector derivada en a sería $(2x, 10y)$ y la superficie unidimensional podría ser expresada parametricamente en términos de α como

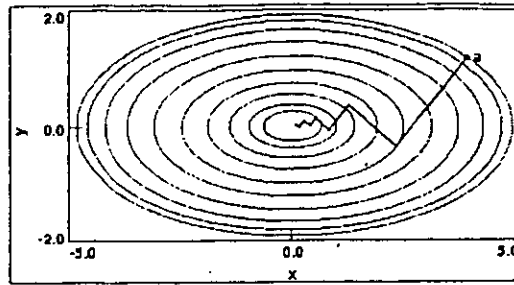


Figura 6-24 Camino seguido en la minimización siguiendo un algoritmo de máximo descenso. Cada una de las 12 iteraciones corresponde a una búsqueda en línea, y fueron necesarias unas 8 evaluaciones de la función por cada línea. Notese cómo las iteraciones no se dirigen directamente al mínimo, haciendo evidente la pérdida de eficiencia discutida en el texto. [Msi97]

$$(x', y') = \left(x_0 + \alpha \left. \frac{\partial \mathcal{E}}{\partial x} \right|_{x_0, y_0}, y_0 + \alpha \left. \frac{\partial \mathcal{E}}{\partial y} \right|_{x_0, y_0} \right) \quad (6.130)$$

donde (x', y') son las coordenadas sobre la línea fuera del punto (x_0, y_0) en la dirección de la derivada en (x_0, y_0) . Si la energía de estos nuevos puntos se calcula y grafica como función de α lo que se obtiene es la curva de la figura (ver figura 6-23).

La búsqueda en línea no depende del método que usamos para encontrar el vector de dirección y se puede elegir de entre muchos métodos para encontrarlo.

Una búsqueda por líneas es eficiente en el sentido de que puede producir un camino corto hasta el mínimo accesible, sobre todo para funciones simples; el problema radica en que es un método costoso en cuanto al número de evaluaciones de la función de energía que hay que hacer. Es necesaria una cantidad relativamente grande de puntos para encontrar con precisión aún el mínimo unidimensional.

Se denomina un *iteración* a un ciclo que concluye cuando se modifica el vector de dirección sobre el que se parametriza.

6.11.2 Máximo descenso

En los métodos de máximo descenso la búsqueda en línea se realiza sobre la dirección del gradiente local $-\nabla \mathcal{E}(x_i, y_i)$. La figura 6-24 muestra el recorrido de minimización seguido según el método de descensos máximos para la función cuadrática simple del ejemplo. Al pasar de una iteración a otra se obtienen rectas mutuamente ortogonales, tal y como ocurría en el caso de la búsqueda en línea original; sin embargo, se observa una oscilación en su convergencia al mínimo. Este tipo de comportamientos es característico de los máximos descensos, sobre todo en superficies de energía con mínimos en "valles" estrechos.

Olvidarnos por completo de la búsqueda en línea y simplemente modificar la posición cada vez que la búsqueda hacia el gradiente de una menor energía tiene la ventaja de que necesitamos evaluar menos veces la función de energía, además de disminuir notablemente la amplitud de las oscilaciones en la trayectoria hacia el mínimo. El resultado de tal minimización es como el de la figura 6-25 en donde el punto de inicio es el mismo que el de la figura 6-24. Cada búsqueda en línea por este método requiere, a lo mas, dos evaluaciones de la función. De la figura 6-24 podemos ver que los pasos hacia el mínimo son mas erráticos, pero este se encuentra en un número muy parecido de iteraciones. El punto crítico es que al evitar búsquedas en línea exhaustivas el número de evaluaciones de la función es solamente entre el 10% y el 20% de lo que se necesitaba por un método puro de búsqueda en línea.

La dependencia exclusiva del máximo descenso en los gradientes es a la vez su fuerza y su debilidad. La convergencia es lenta una vez cerca del mínimo porque el gradiente es casi cero, pero el método es muy fuerte, funcionando inclusive para sistemas muy complejos (y lejos de lo armónico). Es el método que con mayor probabilidad encuentra la estructura de mínima energía, sin importar cual sea la función o el punto de inicio. Por lo tanto máximo descenso es usado comunmente cuando los gradientes son grandes y las configuraciones estan lejos del mínimo. Este es el caso de sistemas cristalinos mal relajados o para moléculas pobremente construidas. Es tan efectivo cuando el gradiente es grande que métodos más avanzados lo utilizan como primeros pasos.

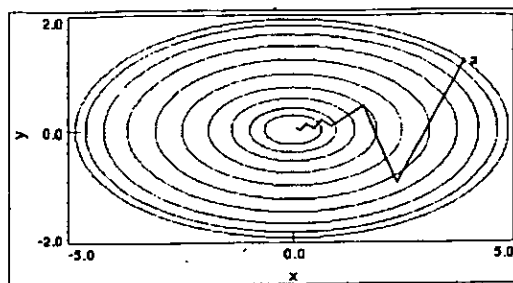


Figura 6-25 Camino seguido por una minimización por máximos decensos sin búsqueda en línea. Converge al mínimo en 12 iteraciones, pero cada iteración requiere sólo 1.3 evaluaciones en promedio, mejorando notablemente el tiempo. [Msi97]

6.11.3 Gradiente conjugado

En la figura 6-25 podemos ver una razón por la cual el método de máximo descenso no converge rápidamente al mínimo. Cada paso debe deshacer algo de lo recorrido por el paso anterior. Incluso en la figura 6-24 todas las líneas están algo desviadas del mínimo ideal, y son necesarias muchas modificaciones a la dirección para llegar suficientemente cerca. Como cada línea debe ser ortogonal a la anterior son necesarias aún más líneas. Lo que observamos, entonces, es un camino de líneas que oscilan cada una corrigiendo un poco del error de la anterior y dando lugar a un nuevo error a corregir.

Por supuesto que preferiríamos tener un método en que un paso no deshiciera lo que se hizo en el paso anterior. Lo que se necesita es un algoritmo en el que en cada paso genere un conjunto completo de direcciones conjugadas, de tal manera que a cada paso se refine la dirección hacia el mínimo. Si este conjunto de direcciones conjugadas cubre por completo el espacio de energías, la minimización a lo largo de cada dirección lleva al mínimo. El método de gradiente conjugado construye y sigue ese conjunto de direcciones.

En gradiente conjugado el vector de dirección \mathbf{h}_{i+1} que sale del punto $i + 1$ se calcula agregando el gradiente \mathbf{g}_{i-1} en el punto $i + 1$ a la dirección anterior \mathbf{h}_i escalada por la constante γ_i así:

$$\mathbf{h}_{i+1} = \mathbf{g}_{i+1} + \gamma_i \mathbf{h}_i; \quad (6.131)$$

γ_i es un escalar que puede definirse de dos maneras. Según el método de Polak-Ribiere γ_i se define como

$$\gamma_i \equiv \frac{\mathbf{g}_{i+1} \cdot \mathbf{g}_{i+1}}{\mathbf{g}_i \cdot \mathbf{g}_i}. \quad (6.132)$$

En el método de Fletcher-Reeves γ_i se define como [Fle80].

$$\gamma_i \equiv \frac{(\mathbf{g}_{i+1} - \mathbf{g}_i) \cdot \mathbf{g}_{i+1}}{\mathbf{g}_i \cdot \mathbf{g}_i}. \quad (6.133)$$

La dirección se usa en lugar del gradiente de la ecuación 6.130 y se comienza otra búsqueda en línea. Una propiedad curiosa de este procedimiento es que el siguiente gradiente \mathbf{g}_{i+1} es ortogonal a todos los gradientes anteriores $\mathbf{g}_0, \mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_i$ y que la dirección \mathbf{h}_{i+1} es conjugada a todas las direcciones anteriores $\mathbf{h}_0, \mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_i$. El nombre de *gradiente conjugado* es en realidad algo equivocado, el algoritmo produce un conjunto de gradientes mutuamente ortogonales y un conjunto de direcciones mutuamente conjugadas.

El método de gradiente conjugado es elegido para sistemas grandes porque a diferencia del de Newton-Raphson (ver sección 6.11.4) en donde se usa una matriz de segundas derivadas, sólo son necesarios los $3N$ gradientes y direcciones anteriores. Sin embargo, para asegurarnos que las direcciones son mutuamente conjugadas, son necesarias algunas minimizaciones por búsqueda en línea en cada dirección. Como estas búsquedas en línea requieren un gran número de evaluaciones de la función de energía por búsqueda, el tiempo por iteración es mayor para gradiente conjugado que para máximo descenso. Esta debilidad es más que compensada por la mucho mejor convergencia.

6.11.4 Métodos de Newton-Raphson

Existen diferentes modelos de Newton-Raphson, entre ellos el Newton-Raphson iterativo, el quasi Newton-Raphson y el Newton-Raphson truncado. Todas estas variantes pueden caracterizarse en forma general por el algoritmo:

1. Dar una propuesta inicial de r_0 .
2. Probar la convergencia.
3. Calcular un Hessiano A aproximado que es positivo definido (el Hessiano es la matriz de segundas derivadas de la energía con respecto a las coordenadas r_0).
4. Resolver para la dirección de búsqueda p_k de manera que

$$\|A_k p_k + g_k\| < \Phi_k \|g_k\| \quad (6.134)$$

en donde Φ_k es alguna cantidad que controla la precisión del p_k calculado.

5. Calcular el largo de paso λ_k apropiado, tal que la energía disminuya una cantidad suficiente.
6. Incrementar las coordenadas:

$$x_{k+1} = x_k + \lambda_k p_k \quad (6.135)$$

7. Volver al paso 2.

Las diferencias entre los diferentes métodos de Newton son

- La forma en que se calcula A y cómo se preserva su carácter de positiva definida durante la minimización.
- La forma y precisión con que se determina la dirección de la búsqueda (i.e. Φ_k).
- Cómo se lleva a cabo la búsqueda en línea y que tan exacta es.

6.11.5 Recocido y templado

Supongamos que un conjunto de N partículas tiene muchas configuraciones de equilibrio posibles. La energía de estas configuraciones es, por lo general, diferente y alguna de ellas será la óptima. Todas ellas corresponden a mínimos locales y están separadas unas de otras por barreras de energía.

Encontrar la estructura óptima siguiendo procedimientos basados en las técnicas tradicionales de minimización requiere de algunos trucos, porque estos métodos normalmente no permiten superar las barreras de energía y llevan al mínimo local. Hay que probar diferentes condiciones iniciales, con la esperanza de que correspondan a diferentes "regiones de atracción" en el espacio de energías, perdiendo mucho tiempo en relajarlas todas ellas. La estructura óptima será aquella con menor energía, pero sólo será óptima dentro de las que encontramos.

La temperatura en dinámica molecular (o en Monte Carlo) nos provee de una forma de "saltar" sobre las barreras de energía. Estados con una energía \mathcal{E} son visitados con una probabilidad $e^{-\mathcal{E}/k_B T}$. Si T es suficientemente grande, el sistema podrá distinguir la existencia simultánea de diferentes mínimos, y pasar más tiempo cerca de los más profundos. Al disminuir lentamente la temperatura T a 0, hay una buena probabilidad de que el sistema pueda elegir el mejor mínimo y bajar en él. Esta consideración es la base de los métodos de *recocido simulado*, en donde el sistema se equilibra a una cierta temperatura y es entonces enfriado lentamente hasta llegar a $T = 0$. Este procedimiento no garantiza que habremos llegado al verdadero mínimo absoluto (dado que no se han hecho suposiciones *a priori* de la estructura óptima), aunque en muchas ocasiones nos lleva a estructuras difíciles de encontrar por intuición y otros métodos.

Se usa comúnmente para optimizar estructuras atómicas, pero su validez es más general: dada una función $Z(\alpha_1, \alpha_2, \dots, \alpha_n)$ de N parámetros, se puede considerar cada uno de estos como un grado de libertad, asignándole una "masa" y mover el sistema con un algoritmo de dinámica molecular o de Monte Carlo para realizar el recocido simulado. Una de las primeras aplicaciones de este método puede encontrarse en el artículo [Kir83] que discute su aplicación al problema del "vendedor viajero".

“Cálculos en la computadora más rápida, ejecutando el algoritmo de minimización y método de integración más eficiente pueden ser inútiles si el sistema se prepara de forma incorrecta.” [Msi97].

Hay muchos factores involucrados en diseñar una estrategia apropiada para llevar a cabo un proceso de minimización en sistema, algunos de los cuales tienen que ver con los algoritmos de minimización. Uno de los pasos más importantes en una simulación es preparar el sistema a ser simulado. El error de preparación puede ser tan simple como un átomo colocado en un mal lugar. Desafortunadamente, no es posible dar una receta única y exitosa para un modelo, ya que mucho depende de los objetivos y expectativas de cada cálculo. ¿Esperamos configuraciones muy lejanas a las de equilibrio?, ¿vamos a comparar las energías en forma cuantitativa?, ¿que hipótesis estamos probando?, etc. Los efectos de amarres (sección 5.3.2), radios de corte (sección 6.3), etc. en los resultados pueden ser evaluados sólo por experimentos preliminares controlados.

A continuación presento algunos de los aspectos a considerar al decidir una estrategia de minimización.

Importancia de una estructura de mínima energía

Es importante tener en cuenta el significado de la configuración de mínima energía. Considerar el valor exacto de la energía es un tanto complicado, porque el valor de energía cero es arbitrario (y probablemente distinto) en cada potencial elegido, por lo que las comparaciones pierden sentido. Sin embargo, tiene sentido hacerlas para diferentes configuraciones de sistemas con el mismo potencial, o cuando se tiene un potencial con un valor cero específico. En principio la energía de una configuración mínima es la entalpía clásica a temperatura absoluta cero, ignorando los efectos cuánticos (en particular la energía de punto cero). Hagler [Hag79] mostraron la forma de calcular la energía de punto cero y libre a mayores temperaturas.

Usando algoritmos diferentes

La elección de cual algoritmo usar depende de dos factores, el tamaño del sistema y el estado actual de la optimización. Mientras las derivadas no se encuentren por debajo de $100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ [Msi97], es probable que el mínimo este aún suficientemente lejos de un mínimo como para que las superficies de energía sean muy lejanas de las armónicas (cuadráticas). Usar algoritmos que asumen la superficie como una cuadrática (Newton-Raphson, quasi-Newton-Raphson o gradiente conjugado) pueden ser inestables en ese punto. El algoritmo de Newton-Raphson es particularmente sensible porque requiere la inversión de la matriz del Hessiano. Por lo tanto, es mejor usar el método de máximo descenso en los primeros 10-100 pasos. Después los métodos de gradiente conjugado y Newton-Raphson pueden usarse para completar la convergencia al mínimo. Para estructuras muy desordenadas podemos encontrar estados metaestables.

Vale la pena comentar algunos aspectos del método de gradiente conjugado: primero, el algoritmo de gradiente conjugado requiere convergencia sobre cada línea, antes de pasar a la siguiente dirección. El gradiente en el paso $i + 1$ debe ser perpendicular a \mathbf{h}_i o la derivación que garantiza un conjunto conjugado de direcciones deja de tener sentido. Segundo, para comenzar los gradientes conjugados, debe elegirse una dirección inicial \mathbf{h}_0 que sea igual al gradiente inicial. Finalmente, es necesario espacio para almacenar un vector extra con los N componentes del gradiente anterior. En un espacio usual son $3N$ derivadas de la energía con respecto a las coordenadas x , y y z de cada átomo. Esto hace gradiente conjugado un excelente método para sistemas que son demasiado grandes para guardar y manipular una matriz de derivadas de segundo orden, como requieren algoritmos como el de Newton-Raphson.

Para sistemas lejos del mínimo en circunstancias muy distintas al armónico, el gradiente conjugado puede minimizar sobre direcciones conjugadas sin converger a un mínimo. Para una descripción detallada del algoritmo se puede acudir a los textos de Press [Pre86] y de Fletcher [Fle80].

Cuando usar restricciones y/o constricciones

En la sección 5.3.2 describí las diferencias entre constricciones y restricciones, así como diferentes casos en que vale la pena usar unas u otras. Específicamente en el ámbito de la minimización es común usar restricciones y constricciones, por ejemplo, si se sabe que se ha de formar una estructura en especial, forzar la atracción entre los átomos participantes. De esta forma eliminamos algo de la dependencia en la configuración inicial. La restricción se elimina (apaga) en algún punto para asegurarnos de encontrar el verdadero mínimo.

Otro ejemplo del uso de restricciones es para modelar sistemas incompletos. Muchas veces es difícil o imposible construir el sistema completo, y debemos restringir algunos átomos en sus posiciones iniciales cristalinas, de lo contrario se moverían. Este tipo de restricciones por lo general son permanentes.

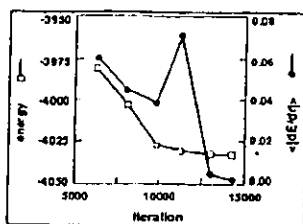


Figura 6-26 Proceso de minimización usando gradiente conjugado. El eje izquierdo muestra la energía como función de la iteración y el derecho el valor absoluto de la derivada. Notese lo mucho que la energía puede cambiar aún para derivadas aparentemente pequeñas. [Msi97]

Para poder realizar minimizaciones útiles es común ir relajando las restricciones y constricciones poco a poco, hasta que el mínimo final representa una configuración no perturbada. En ocasiones es recomendable relajar primero la parte restringida y luego extender la relajación al resto del sistema, aunque también puede ser recomendable el orden opuesto.

Criterios de convergencia

Hay una gran cantidad de criterios para juzgar la convergencia al mínimo. Matemáticamente el mínimo se define como el punto en que las derivadas de la energía son cero, y la matriz de la segunda derivada es positiva definida. Minimizadores que no utilizan el gradiente pueden usar solamente el incremento en la energía y/o criterios de coordinación. En algoritmos que usen el gradiente, las derivadas pueden utilizarse para evaluar la convergencia.

Durante una minimización las derivadas atómicas del conjunto pueden calcularse usando un promedio, una rms, el máximo valor o alguna otra función. El promedio debe ser de los valores absolutos de las derivadas, porque la distribución de derivadas debe ser simétrica respecto al cero. Una derivada rms es una mejor medida que el promedio porque da más peso a las derivadas más grandes, lo que reduce la posibilidad de tener algunas grandes derivadas. Es importante verificar que la máxima derivada no sea demasiado grande.

La pregunta a responder es entonces ¿qué valor del promedio o rms de la derivada constituye una convergencia? El valor específico depende del objetivo de la minimización. Si simplemente queremos relajar átomos encimados antes de comenzar una simulación, minimizar a una máxima derivada de $1.0 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ es generalmente suficiente [Msi97]. Sin embargo, para hacer otro tipo de análisis se acostumbra buscar derivadas promedio menores a $0.02 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$. No es evidente el valor que debe tener esta derivada, y la figura 6-26 muestra que aunque la energía disminuya en forma monótona durante la minimización, la derivada no necesariamente lo hace. Este comportamiento es común para superficies energéticas complicadas.

6.12 Controlando la temperatura

Durante una simulación se tiene un sistema "ficticio" en un estado físico que podemos caracterizar con algunas cantidades, como densidad, temperatura, energía y presión.

Una vez definida la configuración inicial del sistema podemos limitarnos a observar su evolución, pero también podemos hacer algo más, podemos mantenernos pendientes de él y permitirnos "meter la cuchara" y modificar durante la simulación algunas de sus propiedades. Por ejemplo, modificar la energía total puede permitirnos llevar al sistema a estados de equilibrio y alcanzar configuraciones de mínima energía o modificar la temperatura, densidad y presión nos permite simular otros ensambles diferentes al microcanónico. En efecto podemos modificar cualquier parámetro del sistema, el chiste radica en modificar aquellos que nos permitan hacer con el sistema algo interesante y que de preferencia tengan algún significado (o explicación) físico(a).

Los sistemas se comportan de maneras muy diferentes dependiendo de su energía total y quisieramos poder modificar esta a lo largo de la simulación y no sólo determinarla al inicio. La parte potencial de la energía no es fácil de modificar porque no es fácil predecir el efecto de una leve modificación en las posiciones, sin embargo la energía cinética es fácilmente accesible y modificable.

Es importante recordar que la limitante para alcanzar el equilibrio a la temperatura objetivo no es el método de escalamiento de las velocidades, sino diferentes factores, entre los que se encuentran la función de energía potencial.

los parámetros, el acoplamiento entre los modos vibracionales, rotacionales y traslacionales, y el tamaño del sistema (sistemas más grandes toman más tiempo en llegar al equilibrio).

6.12.1 Escalamiento directo

La manera más simple de ajustar la temperatura de un sistema es escalando las velocidades atómicas. Según la ecuación 6.68 tenemos

$$\left\langle \sum_{i=1}^N \frac{|p_i|^2}{m_i} \right\rangle = 2 \langle \mathcal{K} \rangle = 3Nk_B T \quad (6.136)$$

así que escalando las velocidades por un factor β escalamos la temperatura por el factor β^2 . Si $\beta > 1$ la temperatura aumenta, si $\beta < 1$ la temperatura disminuye.

Aplicar cada cierto número de pasos de integración (tiempo) el escalamiento de velocidades permite ir modificando la temperatura y energía cinética y con ella la energía total, pero sabemos que la energía cinética no es una constante durante la simulación, y esperamos que modificarla tenga un efecto sobre la energía potencial. Por ejemplo, aumentar la energía cinética puede llevar a aumentar tanto la energía potencial que puede inducir un cambio en la estructura del sistema.

El caso particular $\beta = 0$ implica "frenar" los átomos. Poco después de detenerlos esperamos que la evolución natural del sistema lleve a aumentar la energía cinética nuevamente a cambio de energía potencial (manteniendo la energía total constante). Dependiendo del algoritmo de integración, del tiempo involucrado y de lo lejos que nos encontremos de una configuración de mínima energía la energía cinética aumentará más o menos. Si la configuración coincide con la de mínima energía no importa cuanto tiempo esperemos los átomos no ganarán ninguna velocidad y la temperatura se mantendrá cero.

La figura 6-27 muestra una función de energía potencial hipotética. Permitir algunos intervalos de integración mueve al sistema sobre la recta A, entre los puntos de retorno a y a' . Si el sistema se encuentra en el punto 1, al frenarlo (hacer cero las velocidades) lo colocamos en el punto 2. Repitiendo el proceso (puntos 3 y 4) nos podemos ir acercando a la configuración de equilibrio. Este es un método de optimización muy común y debemos compararlo con aquellos de la sección 6.11.

Utilizar un factor β constante es útil porque sabemos exactamente lo que hace la simulación, pero podemos tener un mejor control si hacemos β una función de diferentes propiedades del sistema.

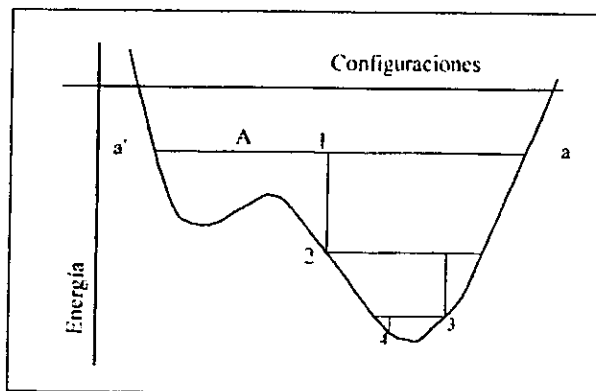


Figura 6-27 En la figura una función de energía potencial hipotética. Permitir algunos intervalos de integración mueve al sistema sobre la recta A, entre los puntos de retorno a y a' . Si el sistema se encuentra en el punto 1, al frenarlo (hacer cero las velocidades) lo colocamos en el punto 2. Repitiendo el proceso (puntos 3 y 4) nos podemos ir acercando a la configuración de equilibrio.

Tal y como hicimos en la sección 5.1.2 para asignar velocidades iniciales para ajustar una cierta temperatura, podemos repetir durante el curso de la simulación, modificando las velocidades según la ecuación 5.8

$$v_{ix}^{nueva}(t) = v_{ix}^{vieja}(t) \sqrt{\frac{\mathcal{K}_{objetivo}}{\mathcal{K}_{actual}}} = v_{ix}^{vieja}(t) \sqrt{\frac{\mathcal{E}_{objetivo} - \mathcal{V}}{\mathcal{E}_{actual} - \mathcal{V}}} = v_{ix}^{vieja}(t) \beta \quad (6.137)$$

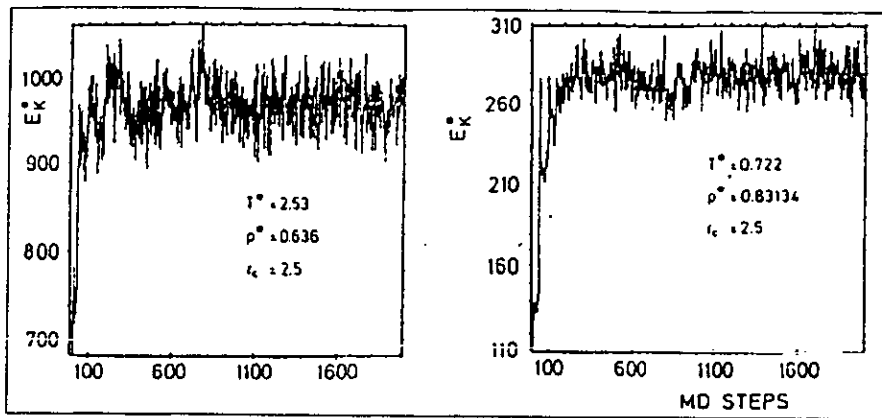


Figura 6-28 Evolución de la energía cinética. Durante los primeros 1000 pasos las velocidades se escalaron cada 50 pasos hasta llegar a las temperaturas deseadas. [Hee90]

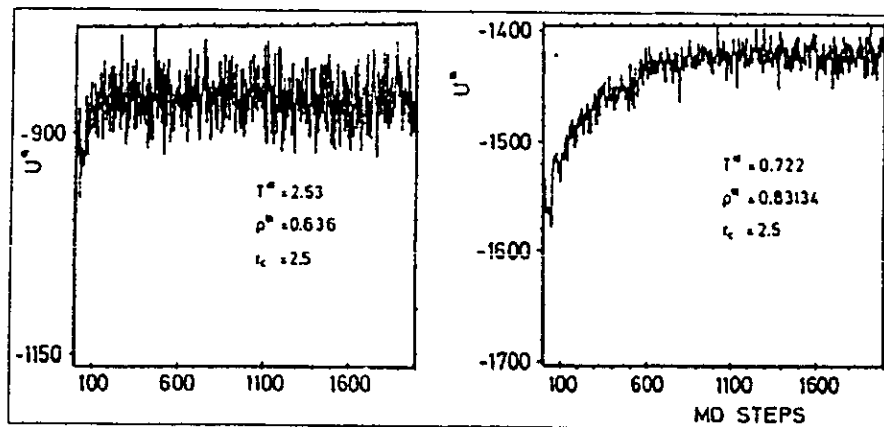


Figura 6-29 Dependencia temporal de la energía potencial. [Hee90]

o la relación

$$v_{ix}^{nueva}(t) = v_{ix}^{vieja}(t) \sqrt{\frac{T_{objetivo}}{T_{actual}}} = v_{ix}^{vieja}(t) \beta \quad (6.138)$$

para asignar una energía cinética o temperatura objetivo. De las ecuaciones 6.137 y 6.138 se deriva la definición útil de β .

Al inicio de la simulación puede ser evidente la temperatura o energías objetivo, pero en el curso de la simulación no necesariamente lo es. Normalmente tiene que calcularse de alguna forma, dependiendo a su vez de otros parámetros, como en el caso de simulaciones de ensamblajes diferentes al microcanónico 6.14.

En el ensamble microcanónico lo usual es encontrar la energía cinética (6.6.2) y potencial (6.6.1) y utilizar la ecuación 6.68 para encontrar la temperatura 6.6.7. para luego utilizar la ecuación 6.138 para escalar las velocidades y llegar al valor adecuado.

La figura 6-28 del Heerman [Hee90] muestra la evolución de la energía cinética para dos sistemas idénticos en diferente estado termodinámico. Se han escalado las velocidades cada 50 pasos de integración. Después de algunos cientos de pasos aparentemente se ha alcanzado la temperatura objetivo $T_{objetivo}$. sin embargo la figura 6-29 muestra la evolución de la energía potencial de los mismos sistemas, en donde podemos ver que la convergencia al equilibrio es más lenta cuando la temperatura es $T = 0.722$. La figura 6-30 muestra la evolución de la energía total, en donde podemos ver que cada escalamiento de las velocidades da lugar a un salto en la energía total. Entre los escalamientos la energía total se mantiene aproximadamente constante, como esperamos. No es absolutamente constante por las aproximaciones en los algoritmos de integración y en el cálculo de la energía potencial al agregar el radio de corte $r_c = 2.5$. Cada vez que una partícula cruza el límite de corte hay una discontinuidad en la energía. Podemos verificar que el sistema se comporta como estamos diciendo al observar la distribución de velocidades con la figura 6-31. Los resultados son buenos aún tomando en cuenta que el sistema sólo cuenta con 256 partículas, y la teoría es mejor para sistemas más grandes.

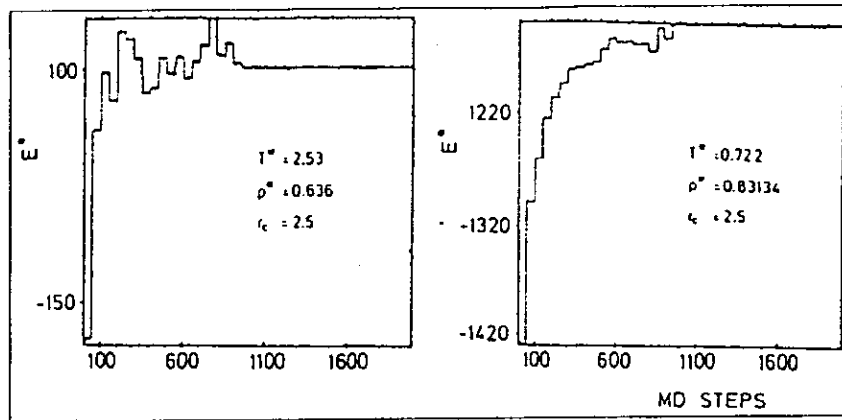


Figura 6-30 Evolución de la energía total. Los saltos en energía corresponden a los escalamientos de las velocidades. [Hee90] A la izquierda de valores instantáneos y a la derecha de promedios.

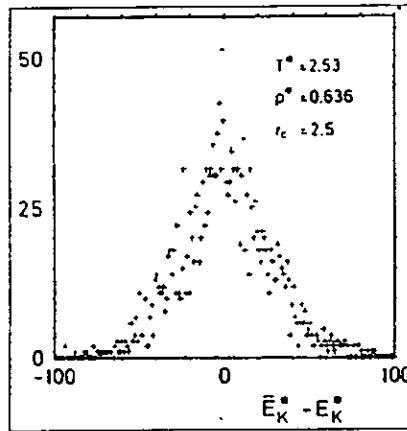


Figura 6-31 Distribución de la energía cinética durante la simulación. [Hee90]

El efecto del radio de corte pequeño elegido es evidente al comparar la energía potencial con radios de corte distintos, cómo se ve en la tabla 8.1 de la sección 8.7.

En la figura 6-30 se puede observar un efecto interesante del ajuste de temperatura. La convergencia puede ser monótona o no. Si la energía cinética oscila (como vemos en la figura 6-28) puede que hayamos elegido hacer el escalamiento en un momento de velocidades particularmente altas o bajas, y llevarnos a una energía total demasiado alta o baja, respectivamente. Si en lugar de la temperatura instantánea utilizamos un promedio en ese intervalo, podemos acercarnos de forma más inteligente al objetivo.

Si estamos en un proceso de optimización para encontrar energía mínima, entonces estaremos disminuyendo la energía, y lo más conveniente es utilizar la configuración de máxima energía cinética (mínima energía potencial) en un intervalo, para hacer sobre ella el escalamiento.

6.12.2 Baño térmico de Berendsen

Una vez equilibrado el sistema, y si no son necesarias grandes modificaciones a la energía cinética (por ejemplo para mantener la temperatura constante), es preferible utilizar un método menos drástico de intercambio de energía entre el sistema y sus alrededores, una buena opción es el método del baño térmico introducido por Berendsen y otros [Ber84]. Cada velocidad se multiplica por un factor λ dado por

$$\lambda = \sqrt{1 + \frac{dt}{\tau}(T_{actual} - T_{objetivo})} \quad (6.139)$$

donde τ es un tiempo de relajación característico. En una buena aproximación, este tratamiento permite mantener la temperatura del ensamble constante, o ajustarla suavemente.

6.12.3 Baño térmico estocástico

Andersen [And80] hace una adaptación diferente de los mecanismos de dinámica molecular para representar la interacción con un baño térmico.

Se selecciona al azar un átomo, y se le ajusta la velocidad según una distribución de Maxwell-Boltzmann (ecuación 5.6). Esto corresponde a una colisión con una partícula de un baño térmico imaginario.

En la descripción original del método, los tiempos entre colisiones con el baño térmico se seleccionan de una distribución de Poisson ($\rho(x) = \sum_{t \leq x} \frac{\lambda^t e^{-\lambda}}{t!}$, $\lambda > 0$), con un tiempo medio entre colisiones específico, pero esto no afecta la distribución en el espacio fase. Si las colisiones se llevan a cabo con poca frecuencia, las fluctuaciones de la energía ocurren lentamente, y las fluctuaciones en la energía cinética (temperatura) serán similares a las de dinámica molecular sin control térmico. Si las colisiones son muy frecuentes las fluctuaciones en la temperatura son dominadas por ellas, y no por la dinámica del sistema. Andersen sugiere un ritmo de colisiones dado por

$$\text{ritmo por partícula} \propto \frac{\lambda_T}{\rho^{1/3} \Lambda^{2/3}} \quad (6.140)$$

donde λ_T es la conductividad térmica. Es interesante que el ritmo disminuya al aumentar el tamaño del sistema.

6.12.4 Método de retroalimentación de Nosé-Hoover

Conocido también como método de *sistema extendido*, me parece particularmente interesante porque las mismas ecuaciones tienen dos diferentes interpretaciones físicas, ambas muy atractivas y bien justificadas. Comienzo con la interpretación de Rapaport [Rap95], y continuo con la de Allen [All87] porque me parece que se complementan perfectamente. Para una exposición detallada es necesario recurrir a los textos completos y a los artículos a los que ellos hacen referencia. Le llamo método de retroalimentación de Nosé-Hoover porque Nosé [Nos84a] [Nos84b] [Nos91] lo propone pero es Hoover [Hoo85] quien hace una reformulación simplificada importante.

Interpretación de Rapaport

Si la temperatura es proporcional al promedio de la energía cinética, la temperatura puede modificarse ajustando el ritmo con el que transcurre el tiempo [Nos84a]. Se introduce una nueva variable s en el Lagrangiano, de manera que sea equivalente a reescalar la unidad de tiempo, además de los términos extra necesarios. Las variables de tiempo son ahora dos, el tiempo *real* o físico t' (de por sí ya en unidades reducidas) y un tiempo escalado (o virtual) t ; la relación entre ambos es a través de su variación

$$dt = s(t') dt' \quad (6.141)$$

El lagrangiano para este sistema extendido es

$$\mathcal{L} = \frac{1}{2} m s^2 \sum_i \dot{\mathbf{r}}_i^2 - \sum_{j>i} V(\mathbf{r}_{ij}) + \frac{1}{2} M_s \dot{s}^2 - n_f T_{\text{objetivo}} \ln s \quad (6.142)$$

donde $n_f = 3N_a + 1$ es el número de grados de libertad (sin descontar aún la conservación del momento lineal). M_s juega el papel de una masa necesaria para construir una ecuación de movimiento para la nueva variable s . El punto significa d/dt . Las ecuaciones de movimiento de lagrange son

$$\ddot{\mathbf{r}}_i = \frac{1}{m s^2} \mathbf{F}_i - \frac{2\dot{s}}{s} \dot{\mathbf{r}}_i \quad (6.143)$$

$$M_s \ddot{s} = m s \sum_i \dot{\mathbf{r}}_i^2 - \frac{n_f T}{s} \quad (6.144)$$

Como la relación entre t y t' depende de la historia completa del sistema tenemos que

$$t = \int s(t') dt' \quad (6.145)$$

por lo que es más conveniente si las ecuaciones se transforman para usar unidades de tiempo físicas [Nos84b] [Hoo85]. Ahora el punto significará d/dt' , y las ecuaciones se reescriben como

$$\ddot{r}_i = \frac{1}{m} \mathbf{F}_i - \frac{\dot{s}}{s} \dot{r}_i \quad (6.146)$$

$$\ddot{s} = \frac{\dot{s}^2}{s} + \frac{m \sum_i \dot{r}_i^2 - n_f T_{\text{objetivo}}}{M_s} \quad (6.147)$$

La primera de estas ecuaciones de movimiento se parece a las convencionales, pero con un término que da la impresión de significar una fricción, pero que sin embargo, puede tener cualquier signo y actuar de forma muy distinta. La segunda ecuación define un efecto de retroalimentación (que Rapaport llama de *feedback*) por medio del cual s varía para regular la temperatura.

Para explicar el origen del término $\ln s$, podemos probar sustituirlo por una función en general $w(s)$. Dado que s es finita, Rapaport afirma que el promedio temporal de \ddot{s} debe ser cero, implicando que

$$m \left\langle \sum_i \frac{\dot{r}_i^2}{s} \right\rangle = \left\langle \frac{dw}{ds} \right\rangle \quad (6.148)$$

El lado izquierdo es simplemente $n_f \left\langle \frac{T_{\text{objetivo}}}{s} \right\rangle$, de tal manera que si igualamos los valores actuales, en lugar de los promedios (integrando), encontramos que $w(s) = n_f T_{\text{objetivo}} \ln s$.

La temperatura no será una constante, pero el efecto de retroalimentación que actúa a través de s garantiza que las fluctuaciones están acotadas y el valor medio es igual a T_{objetivo} .

El hamiltoniano del sistema extendido es

$$\mathcal{H} = \frac{1}{2} m \sum_i \dot{r}_i^2 + \sum_{j>i} \mathcal{V}(r_{ij}) + \frac{1}{2} M_s \left(\frac{\dot{s}}{s}\right)^2 + n_f T_{\text{objetivo}} \ln s \quad (6.149)$$

Se conserva y constituye una manera de probar la calidad de las soluciones numéricas.

Elegir un valor muy grande para M_s resulta en un flujo pequeño de energía entre el sistema y el reservorio, en el límite $M_s \rightarrow \infty$ recuperamos la evolución convencional de dinámica molecular. Por otro lado, si M_s es muy pequeño aparecen oscilaciones de larga duración poco amortiguadas en la energía, resultando en un mal equilibrio. La manera de elegir el valor óptimo para M_s parece ser la *prueba y error*.

Interpretación de Allen

Allen propone que agreguemos al sistema un grado de libertad que represente al reservorio, y simular al sistema completo (de allí a que nombre el método como "de sistema extendido"). Se permite el flujo dinámico de energía entre el reservorio y el resto del sistema, pero le asociamos una cierta *inercia térmica*. Se puede pensar en el método como controlar el volumen de una muestra por medio de un pistón. Nosé [Nos84a] describe la implementación del método. El grado extra de libertad lo denotamos con s nuevamente (para acentuar la similitud con el tratamiento matemático de Rapaport), y su momento conjugado es p_s . Las velocidades de las partículas reales se relacionan con las derivadas temporales de la posición con

$$\mathbf{v} = s \dot{\mathbf{r}} = \frac{\mathbf{p}}{m_s} \quad (6.150)$$

Se asocia la energía potencial extra al término s

$$\mathcal{V}_s = (f + 1) k_B T_{\text{objetivo}} \ln s \quad (6.151)$$

donde f es el número de grados de libertad. La energía cinética también tiene asociado un término

$$\mathcal{K}_s = \frac{1}{2} M_s \dot{s}^2 = \frac{p_s^2}{2M_s} \quad (6.152)$$

donde M_s es el parámetro de inercia térmica con dimensiones [energía · tiempo²] que controla el ritmo de las fluctuaciones de temperatura.

6.12.5 Otros métodos

Existen otros métodos para controlar la temperatura de un sistema, y en particular para mantenerla constante. Quizás los más conocidos son los *Métodos de constricción*, descritos en detalle en los libros de Allen [All87], Rapaport [Rap95] y Heerman [Hee90]. No se basan en un mecanismo de retroalimentación, sino en el uso de constricciones mecánicas, y su justificación no surge del principio variacional de Hamilton, sino del principio de Gauss de mínima constricción [?], que dice que $\sum_i m_i (r_i - F_i/m_i)^2$ es minimizada por el movimiento constreñido. En casos particulares (como el de la dinámica molecular) los resultados son formalmente los mismos que los del principio variacional [Ray72].

6.13 Controlando la presión

Existen diferentes métodos para controlar la presión de un ensamble. El de Berendsen no necesita modificar la forma (ni tamaño) de la celda de simulación, pero sólo es útil para controlar la presión. Los métodos de sistema extendido y de constricción si requieren modificar esta forma, pero a la vez permiten controlar también otros parámetros (que en nuestro caso no necesitamos controlar) cómo la tensión [Msi97].

6.13.1 Método de Berendsen

La técnica que Berendsen y otros [Ber84] describen es particularmente simple para acoplar el sistema a un “reservorio de presión”. Postulan que los cambios en la presión pueden lograrse cambiando las coordenadas de los átomos y el tamaño de la celda unidad (en condiciones de frontera periódicas).

Se añade un término extra a las ecuaciones de movimiento para producir un cambio en la presión, con ello forzamos al sistema a producir un cambio en la presión, haciéndolo obedecer la ecuación

$$\frac{dP}{dt} = \frac{(P_{\text{objetivo}} - P)}{t_p} \quad (6.153)$$

donde t_p es una constante de tiempo que nosotros escogemos. A cada paso de tiempo se escalan las coordenadas de cada átomo por el factor μ

$$r' = \mu r \quad (6.154)$$

y

$$\mu^3 = 1 - \beta_T \frac{dt}{t_p} (P_{\text{objetivo}} - P) \quad (6.155)$$

donde β_T es la compresibilidad isotérmica. Podemos absorber el valor de β_T en el parámetro t_p . Berendsen [Ber84] muestra algunos buenos valores para t_p en el caso de simulaciones de agua.

Este método es fácil de programar y no altera las trayectorias en forma drástica, sin embargo no se ha identificado el ensamble apropiado [All87].

Utilizando un factor idéntico para todas las coordenadas de todos los átomos tiene el efecto de escalar la celda en forma uniforme, pero sin modificar su forma. No siempre es el método adecuado, porque para simular transiciones de fase de cristales puede ser más adecuado modificar tanto la forma cómo el tamaño de la celda [Msi97].

Puede haber grandes fluctuaciones en la presión [Bro84]. Pueden encontrarse presiones negativas porque el virial puede ser negativo (sección 6.6.10), lo que desafía el concepto tradicional de que la presión debe ser positiva. No hay que olvidar el efecto que en el valor de la presión puede tener elegir un radio de corte muy pequeño. Al igual que con la temperatura se pueden aplicar correcciones de largo alcance, pero es necesario tomarlas en cuenta.

6.13.2 Método de Andersen de sistema extendido

Andersen [And80] propuso un método para simulaciones de dinámica molecular de presión constante, que involucra acoplar el sistema a una variable externa V , el volumen de la celda de simulación puede ser visto como representación de la acción de un pistón en un sistema real. El pistón tiene una "inercia" asociada a una masa Q , (con unidades de $[(masa)(longitud)^{-4}]$) y se le asocia la energía cinética

$$\mathcal{K}_V = \frac{1}{2}Q\dot{V}^2. \quad (6.156)$$

La energía potencial asociada a la variable adicional es

$$\mathcal{V}_V = \mathcal{P}_{objetivo}V. \quad (6.157)$$

El lagrangiano es entonces

$$\mathcal{L}_V = \mathcal{K} + \mathcal{K}_V - \mathcal{V} - \mathcal{V}_V. \quad (6.158)$$

Las energías potencial y cinética asociadas a las moléculas pueden reescribirse usando las variables escaladas

$$\mathbf{r} = V^{1/3}\mathbf{s} \quad (6.159)$$

$$\mathbf{v} = V^{1/3}\dot{\mathbf{s}} \quad (6.160)$$

de tal manera que

$$\mathcal{V} = \mathcal{V}(V^{1/3}\mathbf{s}) \quad (6.161)$$

$$\mathcal{K} = \frac{1}{2}m \sum_i v_i^2 = \frac{1}{2}mV^{2/3} \sum_i \dot{s}_i^2 \quad (6.162)$$

y las ecuaciones de movimiento son

$$\ddot{\mathbf{s}} = \frac{\mathbf{f}}{mV^{1/3}} - \frac{2}{3}\dot{\mathbf{s}}\frac{\dot{V}}{V} \quad (6.163)$$

$$\ddot{V} = \frac{(\mathcal{P} - \mathcal{P}_{objetivo})}{Q} \quad (6.164)$$

donde las fuerzas \mathbf{f} y la función de presión \mathcal{P} (ecuación 6.79) se calculan usando coordenadas y momentos normales (no escalados pero si en unidades reducidas). El hamiltoniano del sistema es $\mathcal{H}_V = \mathcal{K} + \mathcal{K}_V + \mathcal{V} + \mathcal{V}_V$ se conserva, siendo igual a la entalpía más un factor $\frac{1}{2}k_B T$ asociado con la energía cinética de las fluctuaciones de volumen [All87]. La conservación de este valor permite realizar pruebas de la calidad de los métodos de integración, que generan trayectorias del ensamble isobarico-isoentálpico (NPH sección 6.14.5), que aún cuando no es uno de los más comunes, si hay quienes han estudiado sus propiedades [Ray81].

Por los detalles de su implementación el mecanismo se adapta naturalmente a los algoritmos de tipo predictor-corrector, y no a los de verlet, aunque hay diferentes modificaciones [Bro84] [Fox84] que solucionan las dificultades.

Igual que en los métodos extendidos de control de temperatura (sección 6.12.4), el parámetro Q representa la "masa del pistón", un parámetro ajustable que determina la interacción entre el sistema original y la extensión. Un valor pequeño da lugar a rápidas oscilaciones del tamaño de la celda, poco amortiguadas por el movimiento de los átomos. Por el contrario, un valor grande conlleva una lenta exploración del espacio de volúmenes; en el límite de una masa infinita recuperamos la dinámica molecular tradicional.

Lo más común es combinar este tipo de métodos con alguno de temperatura constante 6.12 para poder simular un ensamble con NPT constantes 6.14.3. De utilizar un método estocástico las velocidades escaladas se toman proporcionales a $\exp(-mV^{2/3}\dot{s}_i^2/2k_B T)$. Fox y Andersen [Fox84] recomiendan también hacer impactos estocásticos en el pistón, y describen una manera de hacerlo.

6.13.3 Método de Parrinello-Rahman

Este método lo presentaron en detalle en 1981 [Par81]. La diferencia principal con el método de Andersen es que permite cambios en la forma de la celda de simulación, y no sólo en su tamaño. Tiene diversas aplicaciones interesante.

cómo la simulación de sistemas bajo tensiones externas. Para líquidos no es aconsejable porque la celda puede sufrir graves deformaciones, a falta de fuerzas externas de restitución. La excelente adaptabilidad a estudiar transformaciones de fase en sólidos es quizás la principal razón de su amplio uso.

Quien este interesado en aplicar este método puede encontrar la más completa descripción en el artículo de Parrinello y Rahman [Par81] en [Cic87], pero también discusiones al respecto en [All87], [Msi97], [Hee90] y [Rap95].

6.13.4 Otros métodos

Al igual que para controlar la temperatura, para el control de la presión existen otros métodos. Nuevamente los más conocidos son los *Métodos de constricción*, descritos inicialmente por Evans y Morris en [Eva83b] [Eva84], en los que se utiliza el principio de mínima constricción de Gauss [Eva83b] y un multiplicador de Lagrange. Allen [All87] describe con cuidado la teoría y Rapaport [Rap95] y Heerman [Hee90] la implementación.

6.14 Simulando otros ensambles

“El viaje del descubrimiento no busca nuevos parajes sino nuevos ojos.” Marcel Proust.

Integrar las ecuaciones de movimiento de Newton permite explorar una superficie de de energía constante de un ensamble microcanónico. En la sección 6.6 describo las maneras de medir propiedades en este ensamble, mientras que en las secciones 6.12 y 6.13 describo la manera de poder controlar la temperatura y presión respectivamente, ya sea para llevar el sistema a un estado en especial o para mantener constantes estas cantidades.

Dependiendo de cuales variables de estado se mantengan constantes (por ejemplo, la energía \mathcal{E} , el volumen V , la temperatura T , la presión \mathcal{P} o el número de partículas N), se da lugar a diferentes ensambles estadísticos. Una variedad de propiedades estructurales, energéticas y dinámicas pueden ser calculadas de los promedios y/o fluctuaciones de estas cantidades.

Dado que todos los ensambles son representaciones artificiales, dan lugar a promedios que son consistentes unos con otros cuando representan el mismo estado de un sistema (de lo contrario la teoría debería ser modificada o desechada). Sin embargo, las fluctuaciones pueden ser diferentes en diferentes ensamble. Algunas de las fluctuaciones están relacionadas a derivadas termodinámicas, como el calor específico, o la compresibilidad isotérmica. En la práctica, obtener fluctuaciones adecuadas para calcular cantidades puede ser muy complicado, y este método debe ser seguido con cuidado. Allen y Tildesley [All87] discuten las relaciones y reglas de transformación entre diferentes ensambles.

La dinámica molecular se construye de forma más natural en el ensamble microcanónico, pero adaptarla para simular diferentes tipos de ensambles puede ser motivado por interés en estudiar sus propiedades o porque queremos poder realizar una simulación comparable con experimentos realizados bajo condiciones específicas.

En todos los ensambles considerados la cantidad de partículas se mantiene constante.

Los ensambles reciben su nombre de las variables de estado que se mantienen constantes, aunque los más comunes tienen, por lo general, otro nombre histórico.

6.14.1 Ensamble NVE (microcanónico)

Energía y volumen constantes. Se obtiene solucionando directamente las ecuaciones de Newton, sin ningún tipo de control en la temperatura o presión. La energía se conserva aproximadamente, dados los errores de redondeo e integración (sección 6.6.1).

Simulaciones de este tipo no son muy útiles para equilibrar sistemas porque no dan lugar a flujos de energía entre el sistema y sus alrededores, por lo que no es posible colocar al sistema en el estado inicial que nos interesa. Es posible, sin embargo, simular ensambles de este tipo en la fase de recolección de datos explorando una superficie de energía constante y sin las perturbaciones introducidas por los métodos de control de temperatura y presión. En los intervalos entre las recolecciones de datos se puede modificar la energía del ensamble sin mayores dificultades.

6.14.2 Ensamble NVT (canónico)

Temperatura y volúmen constantes. Se obtiene controlando la temperatura, usualmente haciendo un escalamiento directo (sección 6.12.1) mientras se consigue el equilibrio, pasando después a métodos menos drásticos, como de baño térmico (sección 6.12.2 y 6.12.3) o sistema extendido (sección 6.12.4), para la fase de recolección de datos. El volúmen se mantiene constante durante la simulación.

Es el ensamble adecuado cuando lo que se quiere encontrar es la configuración de un sistema periódico. Sin condiciones de frontera periódicas, no quedan definidos claramente el volúmen, la presión y la densidad, por lo que no se pueden llevar a cabo simulaciones a presión constante. Aún con condiciones de frontera periódicas si la presión no es un factor significativo, el ensamble NVT tiene la ventaja de perturbar menos la trayectoria, dada la ausencia de mecanismos de control de presión.

6.14.3 Ensamble NPT (isotérmico-isobárico)

Presión temperatura constantes. Requiere de control sobre la temperatura y la presión. Permite la variación de los vectores de la celda unidad (en sistemas cristalinos), o del tamaño de la celda de simulación en general, cuando se utilizan mecanismos de control de presión que así requieran.

Es la elección adecuada cuando es importante colocar al sistema a una temperatura, presión y densidad específicas. También se utiliza para llevar al sistema a la temperatura y presión objetivo durante la fase de equilibrio, para luego utilizar algoritmos para otros ensambles en la fase de recolección de datos.

Dado que estamos aplicando dos métodos de control, las perturbaciones de la trayectoria pueden ser muy importantes, sobre todo para tiempos largos.

6.14.4 Ensamble NST

Temperatura y tensión constantes. Es una extensión del ensamble de presión constante anterior. Además de la presión hidrostática que se aplica en forma isotrópica, el ensamble de tensión constante permite controlar las componentes xx , yy , zz , xy , yz y zx del tensor de tensión (en ocasiones también conocido como tensor de presión). El ensamble es particularmente útil si a uno le interesa estudiar la relación tensión-deformación de materiales.

6.14.5 Ensamble NPH

Presión y entalpía constantes. Es análogo al ensamble NVE (microcanónico). La suma de \mathcal{E} y PV , es la entalpía H , constante cuando la presión se mantiene fija pero se libera la temperatura.

7. POTENCIALES INTERATÓMICOS

"El electrón no es tan simple como parece" W. L. Bragg [Bri99].

Hace ya mucho tiempo que se conoce la Ley de Coulomb para la interacción entre partículas cargadas, y ha mostrado ser una excelente aproximación, incluso a escalas pequeñas. Tan buena que muchos opinan que describe la interacción real.

¿Para que toda esta filosofía?, quizás por gusto, quizás porque es indispensable mantener siempre a la mano las hipótesis que se han hecho y seguramente porque de las aproximaciones que hagamos dependen los resultados que podemos obtener.

A pesar de que una interacción entre dos partículas que siguen las leyes de Newton interaccionando por medio de la Ley de Coulomb es relativamente simple, cuando se agrega una tercera partícula el problema se vuelve muy complicado. Si el problema se complica mucho al pasar de dos a tres partículas, es de esperarse que para algunos miles sea una verdadera pesadilla. Afortunadamente existe un conjunto de aproximaciones que ha probado simplificar considerablemente el problema, manteniendo aún resultados extremadamente cercanos a la realidad.

Con "realidad" me refiero a que las propiedades calculadas haciendo uso de estas aproximaciones coinciden muy bien con las medidas experimentalmente.

El tipo de aproximaciones que se hacen al formular modelos de interacción las hacemos todos los días, por ejemplo. Si queremos describir la manera en que bolas de billar se mueven sobre una mesa nos basta con usar la mecánica clásica, incluso si queremos tomar en cuenta las energías disipadas por fricción basta con teorías macroscópicas clásicas, no necesitamos calcular lo que le ocurre a cada uno de los electrones y protones de cada uno de los átomos de las diferentes bolas de billar. Tomar en cuenta todos los efectos no es solamente imposible, sino innecesario. Es mucho más sensato aproximar la suma de todas estas interacciones y decir que los objetos se comportan con el modelo de interacción de "bolas de billar", muy simple de entender y de tratar matemáticamente. Es increíble que a pesar de la enormísima cantidad de fenómenos que estamos despreciando, la interacción neta puede ser descrita tan correctamente como queramos por modelos simples para la gran parte de los fenómenos.

Aproximar es, entonces, la esencia de la solución al problema. Por supuesto que dejamos fuera de consideración algunos aspectos, de tal manera que no debemos esperar una concordancia total y completa entre los resultados de la aproximación y los medidos en un sistema real. Hay que saber hasta donde aproximar dependiendo del problema y de la precisión que queremos en el resultado. Quizás el ejemplo más común es el siguiente: Si queremos predecir la órbita de una sonda que vuele entre las lunas de Jupiter, será indispensable conocer con precisión la posición de las principales de estas lunas (sino es que de todas), pero si lo que queremos es calcular el tiempo de vuelo de un tiro con rifle en la Ciudad de Mexico (supondremos que en un campo de tiro, para evitar accidentes), entonces, la posición o incluso la misma existencia de las lunas de Jupiter es despreciable para obtener resultados correctos probablemente en las primeras 10 ó 20 cifras decimales.

La bondad de una aproximación depende de lo que busquemos con ella, por lo que no tiene sentido decir que alguna de ellas es mejor que otra sin especificar el sentido.

Para las primeras simulaciones se eligieron modelos de interacción muy simples, porque la escasa velocidad de cálculo impedía usar modelos más complejos y porque, por el momento, se buscaban los comportamientos generales de las soluciones. Hoy en día somos mucho más ambiciosos en lo que esperamos de las simulaciones, pero también tenemos computadoras mucho más rápidas, y podemos usar modelos notablemente más complejos.

Usar modelos cada vez más complejos puede permitir obtener resultados más cercanos a los experimentales, pero no es el caso. De hecho puede que nos compliquemos la vida innecesariamente incluyendo factores que a la postre resultan sin consecuencias importantes.

Arthur F. Voter [Vot96] hace la observación de que al desarrollarse modelos de interacción más compactos y que a la vez incluyen más física, se aumenta la "precisión por dólar de simulación". Por supuesto esta es una visión un tanto materialista del problema y espero que la preocupación principal de quienes trabajan en el área no sea esa.

Mi intención es presentar una revisión de los principales modelos de interacción, o potenciales interatómicos. Existen muchos más, pero la variedad es tan grande que no se pueden poner todos. Usualmente se clasifican según su forma general (de pares, átomo embebido, etc) o según el tipo de interacción que representan (vg. metálica o iónica, etc.).

En la mecánica clásica (y por lo tanto en la dinámica molecular) el efecto de un potencial sobre las partículas no es directo, sino a través de su gradiente (las fuerzas) y su relación con la derivada del momento. En este sentido la ecuación a solucionar es

$$\mathbf{F}_i = -\nabla V(\mathbf{r}_1, \dots, \mathbf{r}_i, \mathbf{v}_1, \dots, \mathbf{v}_i) \quad (7.1)$$

Podemos decir ahora, que el problema de modelar un material puede sustituirse, para nuestros propósitos, por el problema de encontrar una función de potencial $V(r_1, \dots, r_N)$ para ese material.

7.1 La aproximación de Born-Oppenheimer

“Einstein nunca acepto la mecánica cuántica por el ingrediente de probabilidad e incertidumbre. El decía: Dios no juega a los dados. Parece ser que Einstein estaba doblemente equivocado. Los efectos cuánticos en los agujeros negros no sólo sugieren que Dios si juega a los dados, sino que a veces los tira donde no pueden ser vistos” Steven Hawking.

Si quisieramos considerar todas las interacciones entre los electrones y protones que conforman los átomos tendríamos que tratar con un potencial de esta forma

$$\mathcal{H} = \sum_i \frac{P_i^2}{2M_i} + \sum_n \frac{p_n^2}{2m} + \frac{1}{2} \sum_{i,j} \frac{Z_i Z_j e^2}{|\mathbf{R}_i - \mathbf{R}_j|} + \frac{1}{2} \sum_{n,n'} \frac{e^2}{|\mathbf{r}_n - \mathbf{r}_{n'}|} - \sum_{i,n} \frac{Z_i e^2}{|\mathbf{R}_i - \mathbf{r}_n|} \quad (7.2)$$

en donde los índices i y j se refieren a los núcleos, n y n' a los electrones, \mathbf{R}_i y \mathbf{P}_i son las posiciones y momentos de los núcleos, \mathbf{r}_i y \mathbf{p}_i las posiciones y momentos de los electrones. Z_i el número atómico del núcleo M_i la masa del núcleo y m la masa de los electrones. Esta implícita en esta ecuación la suposición de que todos los protones se encuentran en un núcleo despreciablemente pequeño. Es fácil reconocer las contribuciones de la energía cinética y de las interacciones de Coulomb. En principio lo que tendríamos que hacer es resolver la ecuación de Schrödinger para la función de onda completa $\Psi(\mathbf{R}_i, \mathbf{r}_n)$ y podríamos conocer todo lo referente al sistema.

Por supuesto, tal y cómo dije antes, esto es imposible en la práctica y necesitamos hacer aproximaciones. En 1923 Born y Oppenheimer notaron que los núcleos son mucho más pesados que los electrones, y se mueven en una escala temporal unos dos ordenes de magnitud mayor que los electrones

$$\frac{\omega_{el}}{\omega_{nuc}} \sim \sqrt{\frac{M}{m}} \sim 100 \quad (7.3)$$

por lo que sería sensato suponer los núcleos fijos, mientras nos preocupamos por solucionar la parte electrónica del problema, separando la dependencia en los núcleos $\Xi(\mathbf{R}_i)$ y en los electrones $\Phi(\mathbf{r}_n; \mathbf{R}_i)$, que a su vez depende de las posiciones de los núcleos

$$\Psi(\mathbf{R}_i, \mathbf{r}_n) = \Xi(\mathbf{R}_i) \Phi(\mathbf{r}_n; \mathbf{R}_i) \quad (7.4)$$

Con estas suposiciones el problema lo reformulamos en términos de dos ecuaciones de Schrödinger separadas

$$\mathcal{H}_{el} \Phi(\mathbf{r}_n; \mathbf{R}_i) = \mathcal{V}(\mathbf{R}_i) \Phi(\mathbf{r}_n; \mathbf{R}_i) \quad (7.5)$$

donde

$$\mathcal{H}_{el} = \sum_n \frac{p_n^2}{2m} + \frac{1}{2} \sum_{i,j} \frac{Z_i Z_j e^2}{|\mathbf{R}_i - \mathbf{R}_j|} + \frac{1}{2} \sum_{n,n'} \frac{e^2}{|\mathbf{r}_n - \mathbf{r}_{n'}|} - \sum_{i,n} \frac{Z_i e^2}{|\mathbf{R}_i - \mathbf{r}_n|} \quad (7.6)$$

y

$$\left[\sum_i \frac{P_i^2}{2M_i} + \mathcal{V}(\mathbf{R}_i) \right] \Xi(\mathbf{R}_i) = E \Xi(\mathbf{R}_i). \quad (7.7)$$

La ecuación 7.5 es la ecuación del problema electrónico, considerando los núcleos como fijos. El eigenvalor de la energía $\mathcal{V}(\mathbf{R}_i)$ depende paramétricamente de las coordenadas de los núcleos, y es precisamente esta dependencia a la que llamaremos *potencial interatómico*. Una vez encontrada esta cantidad se integra a la ecuación 7.7, que nos da el movimiento de los núcleos. Cómo esperabamos, en esta ecuación no hay grados de libertad electrónicos, y todos los efectos estan incorporados en el potencial $\mathcal{V}(\mathbf{R}_i)$. Lo acostumbrado es remplazar esta ecuación de Schrödinger con las ecuaciones de Newton, y mover los núcleos en forma clásica. En la sección 4.6.1 dije ya un poco al respecto de las condiciones en que se puede hacer esto de forma segura.

7.2 Suponiendo una interacción más simple

Consideremos primero el caso simple de un sistema de N átomos; la manera más común de dividir la energía total del sistema es tomar en cuenta, primero, la energía de cada átomo como si estuviera solo, luego las energías de cada pareja de átomos, cómo si solo existiera esa pareja, pero tomando en cuenta todas las posibles parejas. Se agregan después las energías de cada triplete de átomos, luego de cuartetos, y así sucesivamente hasta tener un potencial que involucre a todas las partículas a la vez (ver ecuación 7.8).

Según lo anterior la manera de calcular las contribuciones de los subconjuntos a la energía total es

$$V = \sum_i v_1(\mathbf{r}_i) + \sum_i \sum_{j>i} v_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \sum_{j>i} \sum_{k>j} v_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \quad (7.8)$$

en donde he escrito $\sum_i \sum_{j>i}$ para referirme a una suma sobre todos los pares de átomos i y j con $i \neq j$; análogamente para los tripletes y demás. con v_m me refiero a un potencial que involucra m átomos. En la fórmula anterior también he incluido la suposición de que el potencial dependa únicamente de las posiciones de las partículas, y no de otros factores como las velocidades, cargas u otras.

Hay que recordar siempre que si estamos aproximando la energía total por energías de subconjuntos menores, estamos haciendo una hipótesis del tipo de la de superposición. La superposición ha probado ser una manera simple y efectiva de subdividir un problema, la usamos tanto en la gravedad como en el electromagnetismo y otras teorías. Es tan común que ha llegado a tener el rango de principio.

El primer término (v_1) de la ecuación 7.8 representa el efecto de algún campo externo en el sistema. Debe ser un campo relativamente simple para no incluir dependencias en los parámetros que antes mencione, pero aún así puede incluir paredes, gravedad, y otros potenciales interesantes.

El segundo término (v_2) es el potencial de pares y por lo general es el más importante. Depende únicamente de la magnitud de la separación entre los átomos $|\mathbf{r}_i - \mathbf{r}_j| = r_{ij}$, así que por lo general se escribe $v_2(r_{ij})$. De los potenciales de pares hablaremos más en la sección 7.4.

El tercer término, (v_3) involucra los tripletes y ha resultado clave al estudiar líquidos [All87, pág. 8], y de hecho hasta un 10% de la energía de una red de argón puede provenir de estos potenciales. A pesar de que pueden tener una magnitud relativamente importante estos potenciales muy rara vez son utilizados en las simulaciones [Bar71]; la razón es que el tiempo necesario para realizar en una computadora cálculos sobre tripletes es demasiado grande.

El resto de los términos (v_4, v_5 , etc) por lo general son mucho más pequeños que los anteriores, y los despreciamos. Afortunadamente la aproximación de pares es por lo general suficientemente buena, y por lo general es modificada para absorber algunos de los efectos de potenciales de mayor orden, dando lugar al *potencial de pares efectivo*. Usándolo podemos reescribir la ecuación 7.8 como

$$V \approx \sum_i v_1(\mathbf{r}_i) + \sum_i \sum_{i>j} v_2^{eff}(r_{ij}). \quad (7.9)$$

Los potenciales que se usan en simulaciones de computadora por lo general tienen esta forma. Una curiosa, e importante, consecuencia de usar un potencial interatómico efectivo para reproducir datos experimentales es que puede ser diseñado específicamente para ciertas densidades, temperaturas o alguna otra variable. Un potencial de pares simple que no incluya efectos de potenciales de mayor orden no puede justificar su dependencia de otros parámetros.

En los últimos años se han propuesto muchos potenciales que a pesar de no ser tan simples como los de pares, son relativamente fáciles de tratar, y reproducen en forma excelente materiales y propiedades reales.

7.3 Diseñando potenciales

Si vamos a simplificar la interacción coulombiana completa entre todas las partículas de un sistema, por potenciales interatómicos y luego a simplificar estos para suponer que se comportan según un principio de superposición y cómo potenciales de pares, surge la pregunta ¿cómo construimos potenciales para materiales específicos que imiten las propiedades reales?.

Construimos un potencial en dos etapas:

1. Seleccionamos una forma analítica. Esta forma puede incluir dependencias más complicadas que solamente las posiciones atómicas. Puede depender de cantidades geométricas, como las distancias y los ángulos dentro o entre algunas partículas, pero también puede incluir cantidades más generales como el número de coordinación de un átomo.
2. Encontrar una parametrización de las funciones que constituyen la forma analítica que elegimos. Esta parte es muy importante y puede ser en extremo compleja técnicamente.

A lo largo de los años se han probado una gran cantidad de técnicas para llevar a cabo la segunda etapa. En un extremo se encuentran los grupos que sugieren comenzar por una descripción de *primeros principios*, es decir, una descripción en la que la estructura electrónica se toma en cuenta. De esta descripción obtienen una expresión para la energía como función de las posiciones nucleares por medio de aproximaciones sucesivas. En el otro extremo están los grupos que eligen ajustar el potencial a datos experimentales por medio de una parametrización de las funciones analíticas propuestas. El segundo grupo decide dar máxima prioridad al realismo de la descripción, por encima de la relación con los primeros principios.

De cualquier manera, los potenciales se diseñan o construyen teniendo en cuenta un cierto rango de efectividad en mente. Dadas las enormes variaciones en las estructuras electrónicas, puede ser demasiado ambicioso tratar de modelar tanto materiales metálicos, como moléculas diatómicas con el mismo potencial. Inclusive siendo el mismo elemento, el medio ambiente para los átomos puede ser muy diferente.

Si un potencial puede modelar satisfactoriamente diferentes medios decimos que es *transferible*. Si queremos utilizar un potencial para simular, es conveniente tener en cuenta las posibilidades de transferencia y tomar con precaución los resultados obtenidos en condiciones en las que no esté garantizada la validez del potencial. No siempre es evidente que nos encontramos simulando situaciones en las que el potencial pueda no ser válido, pues podemos estar tratando con estados de muy baja coordinación, altas temperaturas o presiones, sin darnos cuenta, para los que el potencial no haya sido construido.

7.4 Potenciales de pares

A continuación presento una serie, de los potenciales de pares más utilizados.

7.4.1 Tipo Coulomb

Incluyo este potencial de todos conocido simplemente porque es la base de todas las teorías de interacción. Es esencialmente el mismo que el potencial gravitacional. Tratar de explicar este parecido en términos de principios más fundamentales es uno de los objetivos de la física. Describe de forma excelente una enorme cantidad de fenómenos. Se expresa mediante

$$V^C(r) = -\frac{kq_1q_2}{r} \quad (7.10)$$

el signo negativo podría ser absorbido por la constante pero usualmente no se hace para mantener implícita la naturaleza atractiva del potencial. Si las cargas q_1 y q_2 son de signos opuestos y repulsiva si son del mismo signo. Para el caso gravitacional las cargas se sustituyen por las masas, se usa una constante de escala diferente y se obtiene siempre un potencial atractivo.

7.4.2 Potenciales Simples

Para trabajar con teorías, o sistemas, muy complicados es necesario usar modelos muy simples, algunos de los cuales son, sorprendentemente exitosos. Escribo a continuación un poco de ellos, mostrando la forma matemática en que se describen y mostrando en la figura 7-1 su forma general.

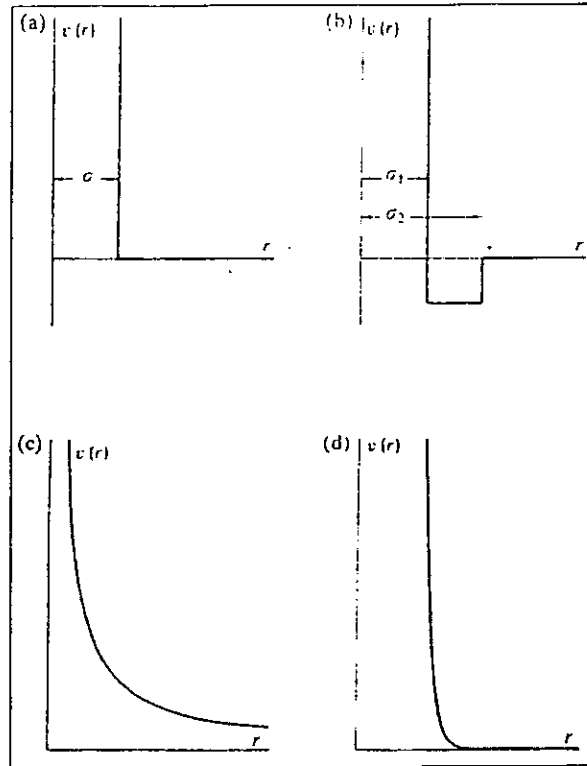


Figura 7-1 Potenciales de pares simples. a) Potencial de esferas duras. b) Potencial de pozo rectangular. c) Potencial de esferas blandas con parámetro $\gamma = 1$. d) Potencial de esferas blandas con parámetro $\gamma = 12$. [All87]

El más simple de ellos es el potencial de esferas duras. Un potencial verdaderamente conveniente para muchos problemas, y que da buenos resultados en simulaciones de sistemas en estado líquido. Corresponde, por ejemplo, al caso que había mencionado antes de las bolas de billar y se describe mediante

$$v^{ED}(r) = \begin{cases} \infty & \text{si } r < \sigma \\ 0 & \text{si } r \geq \sigma \end{cases} \quad (7.11)$$

No tan simple pero incluyendo un pozo en donde tener órbitas cerradas está el potencial de pozo rectangular

$$v^{PR}(r) = \begin{cases} \infty & \text{si } r < \sigma_1 \\ -\epsilon & \text{si } \sigma_1 \leq r < \sigma_2 \\ 0 & \text{si } \sigma_2 \leq r \end{cases} \quad (7.12)$$

que incluye no sólo un parámetro de escala en distancia, sino también un parámetro de escala en energía. El potencial de esferas duras no es más que un caso particular del potencial de pozo rectangular cuando $\epsilon = 0$.

El potencial de esferas blandas no incluye partes atractivas, y su parte repulsiva se puede modificar mediante un parámetro. Lo describimos mediante

$$v^{EB}(r) = \epsilon \left(\frac{\sigma}{r} \right)^\gamma = ar^{-\gamma} \quad (7.13)$$

en donde el parámetro γ por lo general es un entero. Cómo se puede ver en la figura 7-1 al aumentar el valor de γ , el potencial se vuelve más "duro", aproximándose al de esferas duras. A pesar de solo representar efectos repulsivos es curioso que haya casos en que resulte ser la opción más atractiva. Algunos autores le dedican mucho trabajo [Hai92]. Vale la pena hacer notar que si $\gamma = 1$ y $a = -kq_1q_2$ entonces este tipo de potencial se reduce al de Coulomb.

7.4.3 Lennard-Jones

Este potencial es muy común, quizás el más popular en la literatura básica sobre simulación por dinámica molecular, por lo que me gustaría hacer un pequeño repaso simplificado de la manera en que se da lugar a un potencial de este

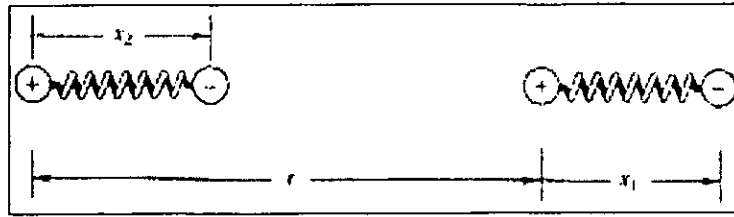


Figura 7-2 Coordenadas de dos osciladores. [Kit86]

tipo.

Interacción atractiva

Consideremos dos átomos de gas noble separados por una distancia r grande en comparación con el radio atómico; en nuestro contexto surge la pregunta ¿cual es la interacción entre ambos?. Si los átomos fueran esferas rígidas con distribuciones de carga fijas, la interacción entre ellos sería nula, porque los potenciales electrostáticos de la nube electrónica y del núcleo se cancelan fuera del átomo. Si fuera así de simple estos átomos vagarían libremente y no se observarían en sus versiones condensadas. Se soluciona el conflicto si suponemos que las cargas electrónicas no están fijas, y que ambos átomos inducen un momento dipolar en el otro, siendo estos momentos dipolares los que dan lugar a una interacción atractiva entre los átomos.

Al afirmar que cada átomo tiene un momento dipolar asociado estamos diciendo que para fines electrostáticos puede ser considerado como una pareja de cargas, una positiva y una negativa separadas por una distancia pequeña. Un modelo simple para poder calcular la interacción es el siguiente: supongamos que lo que tenemos es un par de osciladores armónicos 1 y 2 separados por una distancia r . Cada oscilador tiene una carga $\pm e$ con separaciones x_1 y x_2 , como en la figura 7-2. Las partículas del modelo representan el centro de carga de las nubes electrónicas y de los núcleos atómicos. Las partículas oscilan sobre el eje x , y por simetría se mantienen sobre él siempre. Si denotamos con p_1 y p_2 los momentos, entonces la energía (i.e. el Hamiltoniano) del sistema no perturbado es

$$\mathcal{H}_0 = \frac{1}{2m}p_1^2 + \frac{1}{2}kx_1^2 + \frac{1}{2m}p_2^2 + \frac{1}{2}kx_2^2 \quad (7.14)$$

en donde k es la constante de fuerza asociada al oscilador armónico. Hemos supuesto que, al ser indistinguibles los átomos, tanto las k como las m serán iguales. Cada oscilador armónico tiene por su lado una frecuencia ω_0 , congruente con $k = m\omega_0^2$, (que corresponde con la más fuerte línea de absorción).

Si llamamos \mathcal{H}_0 al hamiltoniano del sistema no perturbado y \mathcal{H}_1 al de la interacción entre ambos osciladores, de la geometría de la figura 7-2 se ve fácilmente que

$$\mathcal{H}_1 = \frac{e^2}{r} + \frac{e^2}{r+x_1-x_2} - \frac{e^2}{r+x_1} - \frac{e^2}{r-x_2}; \quad (7.15)$$

si suponemos que $|x_1| \ll r$ y $|x_2| \ll r$ podemos simplificar 7.15 y obtener

$$\mathcal{H}_1 \simeq -\frac{2e^2x_1x_2}{r^3} \quad (7.16)$$

El hamiltoniano total sería

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1 = \frac{1}{2m}p_1^2 + \frac{1}{2}kx_1^2 + \frac{1}{2m}p_2^2 + \frac{1}{2}kx_2^2 - \frac{2e^2x_1x_2}{r^3}. \quad (7.17)$$

Para resolverlo proponemos la transformación

$$x_s \equiv \frac{1}{\sqrt{2}}(x_1 + x_2) \quad (7.18)$$

$$x_a \equiv \frac{1}{\sqrt{2}}(x_1 - x_2) \quad (7.19)$$

donde los subíndices s y a denotan movimientos simétrico y antisimétrico respectivamente.

Si resolvemos para x_1 y x_2 obtenemos

$$x_1 = \frac{1}{\sqrt{2}}(x_s + x_a) \quad (7.20)$$

$$x_2 = \frac{1}{\sqrt{2}}(x_s - x_a) \quad (7.21)$$

Igualmente se pueden tratar los momentos asociados a ambos modos de oscilación y obtener

$$p_1 = \frac{1}{\sqrt{2}}(p_s + p_a) \quad (7.22)$$

$$p_2 = \frac{1}{\sqrt{2}}(p_s - p_a) \quad (7.23)$$

Después de las transformaciones 7.20 y 7.22 el hamiltoniano total es

$$\mathcal{H} = \left[\frac{1}{2m} p_s^2 + \frac{1}{2} \left(k - \frac{2e^2}{r^3} \right) x_s^2 \right] + \left[\frac{1}{2m} p_a^2 + \frac{1}{2} \left(k + \frac{2e^2}{r^3} \right) x_a^2 \right] \quad (7.24)$$

Por inspección encontramos que las frecuencias de los osciladores acoplados son

$$\omega = \frac{1}{m^2} \left(k \pm \frac{2e^2}{r^3} \right)^{\frac{1}{2}} \simeq \omega_0 \left[1 \pm \frac{1}{2} \left(\frac{2e^2}{kr^3} \right) - \frac{1}{8} \left(\frac{2e^2}{kr^3} \right)^2 + \dots \right] \quad (7.25)$$

$$\text{con } \omega_0 = \sqrt{\frac{k}{m}} \quad (7.26)$$

en donde por supuesto ya esta expandida la raíz como una serie.

La clave viene ahora: la energía de punto cero del sistema es $\frac{1}{2}\hbar(\omega_s + \omega_a)$; debido a la interacción, la suma disminuye del valor desacoplado $\hbar\omega_0$ por

$$\Delta\mathcal{V} = \frac{1}{2}\hbar(\Delta\omega_s + \Delta\omega_a) = -\hbar\omega_0 \frac{1}{8} \left(\frac{2e^2}{kr^3} \right)^2 = -\frac{A}{r^6}. \quad (7.27)$$

Esta diferencia en la energía representa una interacción atractiva que varía cómo r^{-6} modificada por una constante de escala.

La ecuación 7.27 se conoce cómo interacción de van der Waals, interacción de London o interacción dipolo-dipolo inducida. Es la atracción principal en cristales de gases nobles y algunas moléculas orgánicas. Hay que resaltar que la interacción es un efecto cuántico, en el sentido de que si $\hbar \rightarrow 0$, entonces $\Delta\mathcal{V} \rightarrow 0$ (aunque este tipo de límites tiene sus complicaciones filosóficas). La energía de punto cero del sistema disminuye por el acoplamiento de los dipolos. La interacción de van der Waals no depende de la superposición de las nubes electrónicas de ambos átomos.

Para átomos idénticos un valor aproximado para A esta dado por $\hbar\omega_0\alpha^2$, donde $\hbar\omega_0$ es la energía de la línea de absorción más fuerte, y α es la polarizabilidad electrónica [Kit86, cap. 3].

Interacción repulsiva

Hemos supuesto que los centros de carga de los átomos se desplazan formando dipolos, y que la distancia r entre los átomos es mucho mayor que este desplazamiento. Dado que la interacción es atractiva, tiene sentido preguntarnos ¿que pasa si acercamos los átomos?. Conforme se acercan más y más las nubes electrónicas se traslapan (figura 7-3), aumentando la energía electrostática del sistema y dando lugar a una fuerza de separación. Si los átomos se encuentran lejos estas son muy pequeñas, pero pueden llegar a ser muy grandes si la separación es suficientemente pequeña. En buena medida el aumento de esta energía se debe al Principio de Exclusión de Pauli, que Luis de la Peña enuncia cómo: "No existe estado (estacionario de un sistema físico) en el cual dos (o más) fermiones iguales (electrones, protones, neutrones, etc.) compartan los mismos números cuánticos." [Peñ91, pág. 503]. Esto significa que cuando dos átomos idénticos se aproximan y sus nubes electrónicas se traslapan, los electrones de ambos átomos que se encontrarían en el mismo estado cuántico no pueden hacerlo, y algunos de ellos deben ser acendidos a estados superiores libres. Son estados superiores en el sentido de que tienen una mayor energía, que hay que pagar por acercar demasiado los átomos. En la figura 7-4 se muestra un caso a modo de ilustración.

No es fácil calcular la energía de esta interacción fundamentada en primeros principios, pero los datos experimentales se pueden aproximar bien por un potencial empírico de forma $\frac{B}{r^{12}}$ [Kit86, pág. 61], en donde B es una constante positiva para que la interacción sea de naturaleza repulsiva.

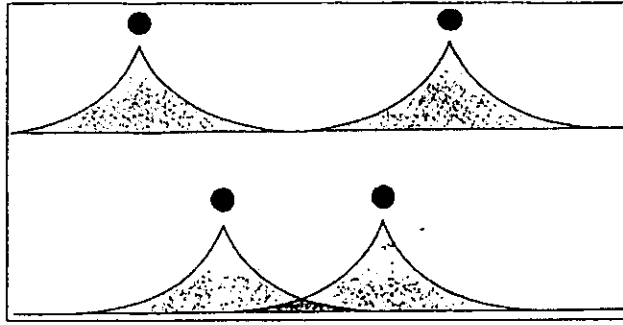


Figura 7-3 Las distribuciones de la carga electrónica de ambos átomos se traslapan al acercarse. El círculo sólido denota el núcleo. [Kit86]

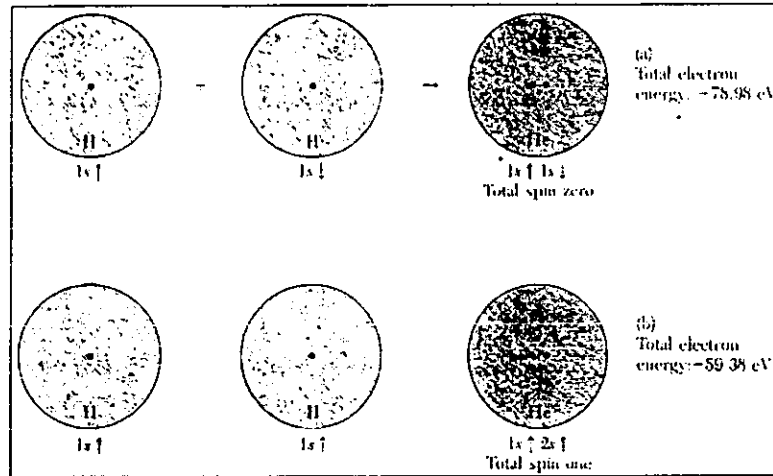


Figura 7-4 Efecto del Principio de Pauli en la energía repulsiva: cómo caso extremo dos átomos de hidrógeno son empujados hasta acercarse tanto que sus núcleos (protones) casi se tocan. La energía de los electrones puede obtenerse de observaciones de átomos de Helio. En (a) los electrones tienen espín antiparalelo y el principio de Pauli no tiene ningún efecto: los electrones están unidos por -78.98 eV. En (b) tienen espín paralelo: el principio de Pauli obliga a los electrones a cambiar de orbital. Los electrones están ahora unidos por -59.38 eV, menor a (a) por 19.60 eV. Esta es la energía que el principio de Pauli ha producido y que tiene un efecto de repulsión. La energía repulsiva coulombiana entre los núcleos ha sido omitida por ser la misma en ambos casos. [Kit86]

Si usamos conjuntamente la parte atractiva resultante de la interacción de van der Waals dipolo-dipolo y la parte repulsiva resultante del principio de Pauli puede ser que se necesite una nueva determinación de las constantes, por ejemplo usando los coeficientes del virial y la viscosidad como parámetros independientes medibles en la fase gaseosa [Kit86]. En este caso lo usual es escribir la energía potencial total como función de la separación r entre los dos átomos así:

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (7.28)$$

donde ϵ y σ son los nuevos parámetros, que se relacionan con los antes dados por $4\epsilon\sigma^6 \equiv A$ y $4\epsilon\sigma^{12} \equiv B$. La ecuación 7.28 se conoce como del potencial de Lennard-Jones, y la figura 7-5 muestra tanto $V(r)$ como $-\frac{dV(r)}{dr} = F(r)$. Los valores de ϵ y σ han de resultar clave y será necesario encontrarlos para tratar con este tipo de potencial. Aunque se usan en cálculos de sólidos, en general provienen de mediciones en gases [Kit86, pág. 63].

Cuando los átomos se acercan mucho también es de esperarse que se incremente la energía electrostática debida a la proximidad de las cargas positivas nucleares, pero por estar estas confinadas al núcleo y las cargas electrónicas distribuidas alrededor, se manifiestan como importantes primero las repulsiones entre los electrones que entre los protones.

Debido a las potencias 6 y 12 para r en la ecuación 7.28 este tipo de potencial se conoce más particularmente como un Lennard Jones 6-12. Es posible aproximar la interacción repulsiva con funciones exponenciales de la forma $\lambda e^{-\frac{r}{\rho}}$, donde ρ es una constante de escala referente al rango de interacción. Las funciones exponenciales son tau

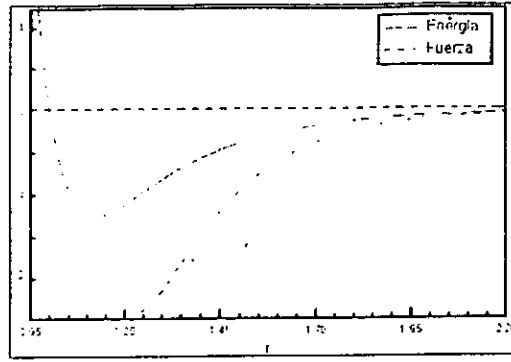


Figura 7-5 Gráfica de la energía y el negativo de la fuerza de Lennard-Jones como función de la distancia entre átomos.

Propiedad	Cu	Ag	Pt	Au	Lennad Jones
$E_c/k_B T_m$	30	28	33	33	13
E_v/E_c	0.33	0.36	0.26	0.25	~ 1
C_{12}/C_{44}	1.5	1.9	3.3	3.7	1

Tabla 7.1 Comparación entre el potencial de Lennard Jones y resultados experimentales.

faciles de tratar analíticamente cómo las de potencias negativas, y pueden llegar a tener algunas ventajas.

El potencial de Lennard Jones no es una buena descripción para cuando existen electrones libres que puedan inducir una atracción o repulsión entre los átomos, pues estas opciones no estaban consideradas en su formulación.

7.4.4 Morse

Al aplicar el potencial de Lennard-Jones para calcular la energía de vibración de moléculas diatómicas se encuentra que los resultados no coinciden aceptablemente con los valores experimentales [Koo90]. Es necesaria una forma analítica del potencial con más parámetros para reproducir simultáneamente la profundidad y localización del mínimo, la frecuencia de vibraciones de pequeña amplitud y el número total de estados ligados. Un potencial que cumple es

$$V(r) = V_0 \left[(1 - e^{-\beta(r-r_{\min})})^2 - 1 \right] \quad (7.29)$$

conocido como *potencial de Morse*. Los parámetros V_0 , β y r_{\min} usualmente se ajustan para dar los valores adecuados de las energías de vibración [Koo90].

7.5 Los problemas con los potenciales de pares

En la sección 7.4.3 presente el potencial de Lennard Jones, quizás el potencial de pares más utilizado, aunque se limita a modelar relativamente bien gases raros, en los que no hay electrones libres. Si nos interesamos en materiales más generales, de usos más prácticos, cómo metales o semiconductores, no podemos modelarlos por potenciales de pares. Consideremos por ejemplo los metales nobles, podemos identificar en ellos algunos indicadores de efectos de muchos cuerpos, efectos que por hipótesis los potenciales de pares no pueden reproducir. En la tabla 7.1 hay algunos datos experimentales para diferentes cristales FCC comparados con los valores predichos por un potencial de Lennard Jones. Otros potenciales dan resultados similares.

$E_c/k_B T_m$ es la relación entre la energía de cohesión y la temperatura de fusión. Esta relación es aproximadamente de 30 en metales, y 10 en sistemas que interactúan por medio de potenciales de pares. La diferencia muestra que los metales exhiben una cohesión extra que es menos efectiva que las fuerzas de pares en mantener el sistema en un estado cristalino.

E_v/E_c es la relación entre la energía de formación de vacancias y la energía de cohesión. Esta relación es de entre $\frac{1}{3}$ y $\frac{1}{2}$ en los metales, pero alrededor de 1 en sistemas de pares (1 exactamente si despreciamos relajaciones). Al formar una vacancia en la estructura de un cristal con número de coordinación C , hemos de pagar la energía E_v para disminuir el número de coordinación de C a $C-1$ para C átomos. En contraste E_c es la energía a pagar para disminuir la coordinación de un solo átomo de C a 0. Con un modelo de pares, en el que hay una energía fija asociada a los enlaces entre pares de átomos, estas dos energías son idénticas, pues ambas se deben a la ruptura de C enlaces, pero esto no es lo que ocurre en los metales reales.

C_{12}/C_{44} es la relación entre dos constantes elásticas de un cristal cúbico. Esta relación es exactamente 1 en sistemas de pares; puede demostrarse analíticamente y se conoce con el nombre de relación de Cauchy [Erc97]. Diferencias de 1 son comunes en los metales. El alto valor para el oro se relaciona con su conocida ductilidad y maleabilidad.

Si consideramos semiconductores, las diferencias con los sistemas de pares son aún peores, por ejemplo, el silicio sufre una serie de transformaciones de fase (de la tetrahedral a la β -tin, a la cúbica simple y a la FCC) cuando se somete a presión. Esto indica que las diferencias de energía entre estas estructuras son tan grandes. En otras palabras, la energía de cohesión es casi independiente de la coordinación (que es muy diferente en las distintas estructuras), mientras que un modelo de energía de pares favorece las estructuras con más enlaces.

Por estas y otras deficiencias [Erc88], desde mediados de los años 80 los investigadores han buscado cómo mejorar el realismo incorporando efectos de muchos cuerpos en los potenciales.

7.6 Potenciales de muchos cuerpos para metales

Un avance significativo se logró durante los años 80 cuando se desarrollaron potenciales de muchos cuerpos para metales basados en el concepto de densidad, o de coordinación como la variable clave a considerar.

El factor clave en este tipo de potenciales es que los enlaces se vuelven más débiles conforme la localidad se llena de átomos. De alguna manera esta sería otra consecuencia de principio de Pauli (ver sección 7.4.3). Una función de la energía de cohesión como función de la coordinación no debe ser necesariamente lineal decreciente, como proponen los potenciales de pares, sino que debe exhibir alguna curvatura positiva, decreciendo más rápido cuando la coordinación es baja y decreciendo suavemente cuando la coordinación aumenta.

La parte repulsiva del potencial aún es tratada como en un potencial de pares, aunque pueda tomar diferentes formas analíticas.

La parte atractiva puede ser obtenida cualitativamente estudiando la relación con el formalismo de tight-binding [Car90]. Un argumento poco riguroso sugiere la forma para la energía del átomo i como

$$E_i \sim \sqrt{\sum_j h_{ij}^2} \propto \sqrt{C_i} \quad (7.30)$$

donde $h_{ij} = h(r_{ij}) = \langle i | H | j \rangle$ son las integrales del traslape entre el átomo i y sus vecinos y C es la coordinación del átomo. En el formalismo de tight-binding los orbitales son localizados y estas funciones son cero más allá de un cierto radio de corte.

La clave de la ecuación 7.30 es que la energía no es proporcional a la coordinación como en los modelos de pares, sino a su raíz cuadrada. Este tipo de relación va en la dirección correcta para resolver los problemas mencionados en la sección 7.5 y asociados a la dificultad de los potenciales de pares para tratar con metales. Por ejemplo

$$\frac{E_v}{E_c} \sim \frac{C[\sqrt{C} - \sqrt{C-1}]}{\sqrt{C}} \sim \frac{1}{2} \quad (7.31)$$

y según algunos autores [Erc97, pág. 41] la relación de Cauchy $C_{12} = C_{44}$ ya no se satisface.

Hay varios métodos, similares a este para construir potenciales de muchos cuerpos para metales, todos ellos basados en la forma analítica

$$V = \frac{1}{2} \sum_i \sum_{j>i} \phi(r_{i,j}) + \sum_i U(n_i) \quad (7.32)$$

donde $\phi(r)$ es el potencial de pares y $U(n)$ es una función que da la energía de un átomo como función de una cierta coordinación generalizada n . Para un átomo i en particular n_i se construye con la superposición de las contribuciones

de sus átomos vecinos, de esta forma

$$n_i = \sum_i \sum_{j>i} \rho(r_{ij}) \quad (7.33)$$

donde $\rho(r)$ es una función de la distancia, usualmente decreciente y de corto alcance. Dentro de este mismo esquema general se encuentran la *teoría de medio efectivo* [Jac87], y los potenciales del *modelo de pegamento* (glue model) [Erc88], del *modelo de átomo embebido* (Embedded atom model) [Daw84] [Bas86] al que trato en detalle en la sección 7.6.1, y los *potenciales de Finnis-Sinclair* [Fin84]. Aún compartiendo la forma analítica los diferentes potenciales difieren notablemente en el procedimiento que usan para construir las funciones $\phi(r)$, $\rho(r)$ y $U(n)$ que constituyen el modelo, con parametrizaciones muy diferentes para un mismo material. Si las mismas formas analíticas con distintas parametrizaciones pueden llevar a resultados tan diferentes como para distintas formas, podemos concluir que ajustar es lo verdaderamente esencial al construir un potencial.

Para algunos grupos la simulación de dinámica molecular es esencialmente una prueba para sus potenciales, por ejemplo evaluando el coeficiente de expansión térmica o el punto de fusión.

Es interesante notar que la ecuación 7.32 es invariante respecto a la transformación

$$\hat{\phi}(r) = \phi(r) + 2\lambda\rho(r) \quad (7.34)$$

$$\hat{U}(n) = U(n) - \lambda n \quad (7.35)$$

para cualquier valor de λ . Esto significa que no existe una elección única para $\phi(r)$ y $U(n)$, y que uno puede arbitrariamente imponer una condición como $U'(n_0) = 0$ (en donde n_0 es, por ejemplo, la coordinación de un átomo del bulto en equilibrio y $T = 0$). También significa que si $U(n)$ es una función lineal de n , entonces todo el esquema se reduce al de un potencial de pares. Los efectos de muchos cuerpos están, entonces, relacionados con la curvatura de $U(n)$. De hecho $U''(n)$ es invariante con respecto a 7.34.

Para implementar estos potenciales en un programa de dinámica molecular se necesita evaluar las fuerzas

$$\mathbf{F}_i = - \sum_i \sum_{j>i} (\phi'(r_{ij}) + [U'(n_i) + U'(n_j)] \rho'(r_{ij})) \frac{\mathbf{r}_{ij}}{r_{ij}}. \quad (7.36)$$

Este cálculo es un tanto más complicado que el requerido para un sistema de pares. Energía y fuerzas aún pueden ser obtenidas usando las distancias de pares como entrada, ya que dada la simetría no hay ninguna dependencia en términos angulares. Por otro lado, la falta de verdaderas fuerzas angulares hace difícil usar estos esquemas para modelar metales en que los efectos covalentes sean importantes, por ejemplo, metales de transición. Se están desarrollando nuevos esquemas más poderosos para tratar con estos materiales [Car90].

7.6.1 Modelo del átomo embebido (EAM)

En el marco del EAM, la energía interna total de un sistema de N átomos es la energía requerida para embeber estos N átomos en el gas homogéneo de electrones causado por los demás átomos, más una corrección por la energía de las interacciones de pares. La energía total \mathcal{E}_{total} la expresamos como ([Daw83] [Daw84])

$$\mathcal{E}_{total} = \sum_i F_i(\rho_i) + \sum_{i>j} \phi(r_{ij}), \quad (7.37)$$

donde $F_i(\rho_i)$ es la energía de embebimiento requerida para colocar el átomo i en una densidad ρ_i , $\phi(r_{ij})$ es el potencial de pares entre los átomos i y j separados por una distancia r_{ij} ; ρ_i denota la densidad electrónica en el átomo i debida a todos los demás átomos. De acuerdo a la aproximación de superposición lineal, esta densidad electrónica puede descomponerse en la suma de densidades electrónicas $f(r_{ij})$ del átomo individual j , es decir

$$\rho_i = \sum_{j \neq i} f(r_{ij}). \quad (7.38)$$

Notese que la función de embebimiento $F(\rho)$ es universal, en el sentido de que no depende de la fuente de densidad electrónica, por lo que la misma función de embebimiento puede ser usada para calcular la energía de un átomo en un compuesto que en un material puro; hay detalles que no trato aquí, pero que pueden ser revisados en [Foi86].

La función 7.37 puede tener diferentes interpretaciones según la teoría que le de lugar [Fin84] [Erc86]. Esto quiere decir que $F(\rho)$, $\phi(r)$ y $f(r)$ pueden tener diferentes forma funcional. Diferentes autores hacen diferentes propuestas.

destacando las de Finnis y Sinclair [Fin84], Voter y Chen [Vot87], Ercolessi, Tosatti y Parrinello [Erc86] y otras derivadas de primeros principios [Jac87] [Daw89a].

Ye y Cai [Ye96] sugieren que en la práctica uno debe preferir una aproximación empírica o semi empírica para obtener estas funciones. En algunos modelos EAM [Foi86] [Vot87] se asume que ϕ tiene una forma analítica, y $F(\rho)$ es una función tabulada de la densidad electrónica. Johnson postula un modelo EAM analítico [Joh88] para cristales FCC, pero Ye y Cai le encuentran diversas deficiencias, por lo que definen sus propias formas analíticas para $F(\rho)$, $\phi(r)$ y $f(r)$, físicamente plausibles y ajustadas de tal manera que reproduzcan propiedades físicas básicas de metales puros.

Banerjea y Smith han mostrado que puede usarse una función exponencial simple para representar bien la densidad electrónica en diversas circunstancias [Ban88]. Como primera aproximación se puede asumir que la densidad electrónica sigue la forma exponencial

$$f(r) = f_e e^{-\chi(r-r_e)}, \quad (7.39)$$

donde f_e es una constante de escala, r_e es la distancia de equilibrio y χ es un parámetro ajustable que debe ser determinado. Al obtener funciones adicionales del EAM, sugieren Ye y Cai que se tenga cuidado de asegurar que la función de embedding tenga curvatura positiva, lo que garantiza que el orden de la fuerza de las uniones disminuya al aumentar la coordinación [Daw89b]. En su modelo toman la función de embebimiento según la forma universal propuesta por Banerjea y Smith [Ban88], y entonces

$$F(\rho) = -F_0 \left[1 - \ln \left(\frac{\rho}{\rho_e} \right)^n \right] \left(\frac{\rho}{\rho_e} \right)^n + F_1 \left(\frac{\rho}{\rho_e} \right). \quad (7.40)$$

donde el segundo término de la izquierda es lineal y ha sido agregado a la fórmula original de Banerjea y Smith, que según Ye y Cai no modifica la curvatura positiva de la función.

Definen el potencial de pares con una forma analítica como la del modelo de Rose y otros [Ros84]

$$\phi(r) = -\alpha \left[1 + \beta \left(\frac{r}{r_a} - 1 \right) \right] e^{-\beta \left(\frac{r}{r_a} - 1 \right)} \quad (7.41)$$

Que según ellos da lugar a un conjunto de funciones para el EAM que describe las propiedades lejos del equilibrio.

En las ecuaciones 7.40 y 7.41 ρ_e representa la densidad electrónica en el equilibrio y F_0 y n son dos constantes. Ye y Cai toman $F_0 = \mathcal{E}_c - \mathcal{E}_v^f$ y $n = 0.5$, donde \mathcal{E}_c y \mathcal{E}_v^f son las energías de cohesión y de formación de vacancias, respectivamente. No dan ninguna justificación física para esto, y argumentan que es simplemente por conveniencia. El resto de los parámetros (r_a , F_1 , α , β y χ) deben determinarse para sistemas de composición pura. f_e es un factor de escala y se cancela automáticamente debido a la forma de las ecuaciones, por lo que sin pérdida de generalidad se puede considerar $f_e = 1$. Para compuestos no necesariamente es tan fácil, y f_e debe especificarse.

Ye y Cai explican la manera de encontrar valores para los parámetros para siete metales diferentes. Las figuras 7-6 muestran las funciones $\phi(r)$ y $F(\rho)$ para los siete metales.

Me gustaría recomendar la explicación que el sueco Mats Carlberg hace del EAM en su tesis de licenciatura [Car96].

7.7 Potenciales de muchos cuerpos para semiconductores

Los semiconductores presentan aún más problemas que los metales. Consideremos por ejemplo el silicio y algunas de sus propiedades [Erc97]

- La fase más estable en ausencia de presión tiene la estructura del diamante: muy abierta y con coordinación 4.
- Al aplicarle presión aparecen nuevas estructuras con creciente coordinación (β -tin, cúbica simple, FCC, etc.), indicando que no son tan lejanas en energía.
- En su fase líquida es un metal, y es más denso que cuando sólido (anómalo).

Con esto basta para garantizar que construir un potencial para Si no es un trabajo simple. De cualquier manera, dada su importancia tecnológica hay muchos grupos luchando en esa dirección, con diferentes niveles de éxito.

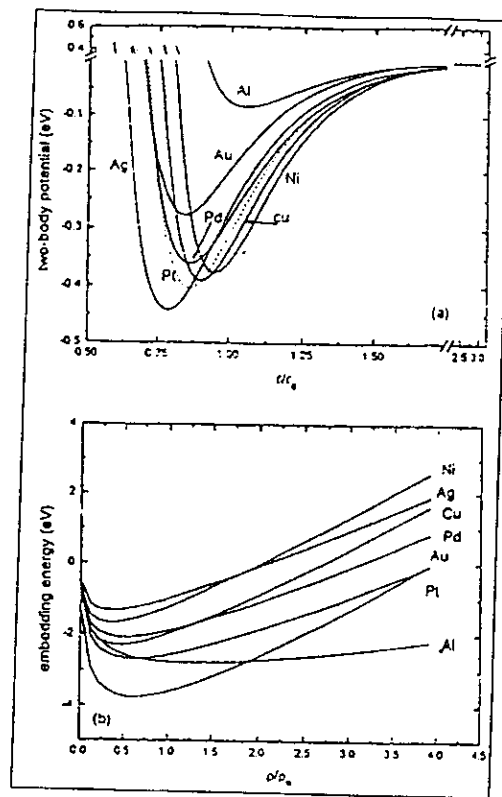


Figura 7-6 Energía de configuración correspondiente al a) potencial de pares y b) energía de embebimiento. [Ye96]

7.7.1 Potencial de Stillinger-Weber

El potencial de Stillinger-Weber [Sti85] es uno de los primeros intentos de modelar un semiconductor con un modelo clásico. Esta basado en el uso de un término de dos y uno de tres cuerpos:

$$V = \frac{1}{2} \sum_{i,j} \phi(r_{ij}) + \sum_{i,j,k} g(r_{ij})g(r_{ik}) \left(\cos \theta_{ijk} + \frac{1}{3} \right)^2 \quad (7.42)$$

donde θ_{ijk} es el ángulo formado por los enlaces ij e ik . $g(r)$ es una función decreciente que se desvanece en su radio de corte localizado en un punto intermedio entre las capas de primeros y segundos vecinos. La intención es favorecer aquellas configuraciones donde $\cos \theta_{ijk} = -\frac{1}{3}$, es decir, los ángulos que corresponden a la estructura tetrahedral del diamante, y hacer de esta estructura compacta la más estable.

Stillinger-Weber es una descripción relativamente realista de los cristales de silicio. Desgraciadamente su origen tetrahedral acarrea problemas en otras situaciones: no puede predecir las energías correctas para las facetas no tetrahedrales encontradas bajo presión. La coordinación del líquido es muy baja. Las estructuras formadas en su superficie no son las correctas.

Tal y cómo se definio antes en la sección 7.3 estos son llamados problemas de transferencia porque el potencial no se adecúa a situaciones diferentes de aquellas para las que se diseño.

Investigaciones recientes [Car90] muestran que para construir modelos más realistas es necesario considerar formas analíticas que tomen en cuenta el concepto de medio *ambiente local* (local environment), y hagan la fuerza de los enlaces dependiente de el.

Sin embargo, potenciales basados en conceptos similares, y que usan solamente cantidades geometricas cómo las distancias y los ángulos de enlace, también son usados hoy en día al simular sistemas organicos, en donde normalmente se les llama *campos de fuerza* (force fields). A pesar de la poca posibilidad de transferir estos potenciales, y de las limitaciones que esto impone a cambiar la topología de una molécula al romperla y reformar sus enlaces, o a trabajar a altas temperaturas y/o presiones, son capaces de modelar con una muy alta precisión las propiedades estructurales y dinámicas de una gran variedad de moléculas, y por ello son muy usados en investigación y en la industria.

7.7.2 Potencial de Tersoff

La familia de potenciales desarrollada por Tersoff [Ter88] se basa en el concepto de *orden de enlace* (bond order): la fuerza de un enlace entre dos átomos no es constante, sino que depende del medio ambiente local. La idea es similar a la del modelo de pegamento, para metales, de usar la coordinación de un átomo como la variable que determina la energía. En semiconductores, sin embargo, el énfasis se hace en los enlaces en lugar de en los átomos, porque es allí donde la carga electrónica tiene su efecto.

A primera vista un potencial de Tersoff da la impresión de ser un potencial de pares:

$$V = \frac{1}{2} \sum_{i,j} \phi_R(r_{ij}) + \frac{1}{2} \sum_{i,j} B_{ij} \phi_A(r_{ij}) \quad (7.43)$$

donde R y A se refieren a las partes repulsiva y atractiva respectivamente. Sin embargo no es un potencial de pares porque B_{ij} no es una constante, es el *orden del enlace* para el enlace que une los átomos i y j , y es una función decreciente de la coordinación G_{ij} asignada al enlace

$$B_{ij} = B(G_{ij}). \quad (7.44)$$

Por su parte G_{ij} se define como

$$G_{ij} = \sum_k f_c(r_{ik}) g(\theta_{jik}) f(r_{ij} - r_{ik}) \quad (7.45)$$

en donde $f_c(r)$, $f(r)$ y $g(\theta)$ son ciertas funciones adecuadas [Erc97]. La idea central es que lo que el enlace ij se debilita esta determinado por donde se encuentran otros enlaces vecinos. Aparentemente se necesitan términos angulares para construir un modelo realista.

Este esquema trabaja en un espectro de situaciones más amplio que el potencial de Stillinger-Weber, sin embargo, no esta exento de problemas. Uno de los principales sería que es difícil de ajustar, por tener 6 funciones libres, además de terminos angulares.

Una aplicación común de estos potenciales son los hidrocarburos [And80].

7.8 Fuerzas de largo alcance y sumas de Ewald

Hasta ahora hemos supuesto que las interacciones son de corto alcance, es decir, que el efecto que el átomo i tiene en la fuerza sobre el átomo j se desvanece (se hace cero) cuando su separación r_{ij} es mayor que un cierto radio de corte R_c . Esto es razonable en muchos casos, pero hay circunstancias en que es necesario tomar en cuenta la influencia de fuerzas de largo alcance. Por ejemplo si queremos hacer cálculos cuando:

- En el medio hay iones cargados, por ejemplo, en una simulación con sales del tipo NaCl. En este caso las interacciones entre iones son de Coulomb y de largo alcance. con un potencial de la forma $\sim \frac{1}{r}$.
- Hay partículas neutrales, pero polarizadas. como en el agua. En este caso habra interacciones de tipo dipolo-dipolo con la forma $\sim \frac{1}{r^3}$.
- Atracciones del tipo de van der Waals entre átomos, como las de la ecuación 7.27 en la sección 7.4.3, resultantes de momentos dipolares inducidos que pueden ser consideradas bajo algunas circunstancias como de largo alcance, dependiendo de las precisiones y las constantes involucradas. Un ejemplo sería la punta de un microscopio de tunelaje, que es atraída a una superficie a decenas de Angstroms de distancia por una atracción de van der Waals.

En estos casos no existe un radio de corte adecuado. Más importante aún es que cada partícula puede interactuar con todas las imagenes de las otras partículas en las celdas imagen (ver sección 6.3), y la regla de imagen mínima introducida en la sección 6.3.7 deja de ser aplicable. Es posible que una partícula interactúe incluso con sus propias imagenes.

Un método estándar para tratar con fuerzas de largo alcance son las sumas de Ewald.

Si simulamos iones cargados, podemos suponer que cada ion está rodeado de una distribución esférica de carga de signo opuesto y que neutraliza al ion. Esta distribución es localizada, es decir, está contenida dentro de una esfera definida por algún radio finito, lo que da lugar a un efecto de pantalla para las interacciones entre iones, las que podemos tratar con las técnicas para potenciales de corto alcance.

Para recuperar el sistema original, consideramos las interacciones coulombianas de distribuciones de carga similares centradas en los iones, pero del mismo tipo (mismo signo) que los iones originales. Estas distribuciones cancelan exactamente a las anteriores, de tal manera que las interacciones totales obtenidas serán las del sistema original formado de cargas puntuales únicamente. La interacción de las distribuciones que se cancelan se calcula en el espacio recíproco. Las transformaciones de Fourier necesarias son particularmente simples cuando la forma de las distribuciones se elige arbitrariamente como una gaussiana, para lo cual no hay ningún impedimento (sección 6.9 y [All7, sección 5.5]).

7.9 ¿Realmente son indispensables los potenciales?

En 1985 Car y Parinello [Car85] desarrollaron un poderoso método que permitía realizar simulaciones de dinámica molecular en donde los iones se movían en forma clásica, pero por la acción de fuerzas obtenidas al solucionar el problema de la estructura electrónica. De esta manera habían saltado la necesidad de potenciales empíricos a cambio de un costo enorme en poder de cómputo. Este importante desarrollo detona una de las más grandes líneas de investigación hoy en día dentro de la física de materia condensada, la conocida como *dinámica molecular de primeros principios* (FPMD por sus siglas en inglés), *dinámica molecular ab-initio* o simplemente método de Car-Parinello.

Al aparecer los métodos de simulación ab-initio con sus ventajas en precisión aparece también en el mapa la pregunta ¿significa esto la desaparición de las simulaciones por dinámica molecular clásica y la muerte de los potenciales empíricos?. Según Ercolessi [Erc97] (y yo estoy de acuerdo) la respuesta es no.

Mientras que actualmente es posible realizar simulaciones de dinámica molecular de sistemas con millones de partículas en las computadoras más poderosas, o simular tiempos del orden de nanosegundos ($10^{-9}s$), el límite actual para las simulaciones ab-initio es del orden de unos miles de átomos y sus tiempos de simulación normalmente se miden en picosegundos ($10^{-12}s$) (ver sección 4.6). Esto significa que problemas que requieran grandes tamaños y/o tiempos largos sólo pueden ser atacados sólo por métodos clásicos.

Aunque la velocidad de las computadoras se ha incrementado de una manera asombrosa, también las dimensiones de los problemas que nos interesa resolver, y lo más probable es que la mayoría de las técnicas de simulación que conocemos hoy en día se mantengan en uso por algún tiempo, incluso que veamos aparecer nuevas técnicas para simulación clásica.

Aún los sistemas más grandes y mayores tiempos de simulación que se logran con los más poderosos equipos de cómputo son minúsculos comparados con los sistemas reales estudiados. Siempre podemos preguntarnos algo que no podremos contestarnos con las computadoras más potentes, es decir, nunca habrá poder de cómputo suficiente, lo que a su vez significa que siempre habrá lugar para simplificaciones sensatas.

7.10 Datos ab-initio por ajuste de fuerzas

En la sección 7.7.1 vimos como, en nuestra búsqueda de realismo al modelar un potencial, la forma analítica puede volverse sumamente complicada. Parametrizar la función que aparece en un Hamiltoniano complejo puede volverse un trabajo muy pesado, por lo que la gente normalmente recurre a parametrizaciones simples. Por ejemplo, los potenciales de Tersoff (ver sección 7.7.2) tienen solamente 12 parámetros, un promedio de sólo 2 parámetros para describir cada una de las 6 funciones que constituyen el modelo.

Un programa de dinámica molecular de primeros principios puede fácilmente producir un enorme número de cantidades útiles y datos confiables, por ejemplo las fuerzas sobre los átomos en una variedad de geometrías y coordinaciones diferentes. Estos datos pueden complementar las cantidades experimentales normalmente usadas para ajustar los potenciales, ayudando a construir potenciales más realistas. Recientemente se ha desarrollado el método de *ajuste de fuerzas* [Erc94], que consiste en un proceso de optimización numérica que trata de ajustar tan cercanamente como sea posible las fuerzas derivadas de simulaciones ab-initio a un potencial clásico, usando una

parametrización rica (entre 10 y 15 parámetros para cada función del potencial). Al incluir explícitamente diferentes geometrías y diferentes temperaturas en el conjunto de datos, podemos tratar el problema de la transferencia del potencial en sus raíces. Según Ercolessi podemos decir que “el potencial se construye aprendiendo de primeros principios”.

Para el estudio de cúmulos metálicos y covalentes es necesaria una descripción más compleja, porque interactúan por medio de fuerzas que dependen del ángulo y que no son aditivas. Para el trabajo con materiales grandes (muchos átomos) son aplicables los métodos anteriores de ajuste de parámetros, sin embargo es necesario repetir el trabajo con algunas modificaciones para poder aplicarlo a cúmulos porque el menor grado de coordinación de los átomos dada la mucho mayor relación superficie/volumen modifica la respuesta [Gar98]

Hay varios grupos trabajando en este frente, de tal manera que podemos esperar que en los próximos años aparezcan potenciales para clases de materiales hasta ahora inaccesibles a la simulación de dinámica molecular.

Parte III

El banco de trabajo

8. QUEREMOS UN BUEN SIMULADOR

“El Apollo 11 llegó a la luna con dos computadoras del tamaño de una caja de zapatos, 64 kb y el poder de cómputo de una calculadora moderna, así que la programación inteligente es más importante que la fuerza bruta”.

8.1 ¿Que es bueno?

En la dinámica molecular, al igual que en la mayoría de los ámbitos, no es fácil definir que es bueno y que es malo. En lugar de meterme en una discusión filosófica me parece pertinente separar las bondades de un programa de simulación en sus bondades como programa y sus bondades para la física. De cada una de las partes menciono algunas de las propiedades a considerar.

8.1.1 Bondad para la física

Dentro de las propiedades importantes para poder considerar como buena (o mala) una simulación, desde el punto de vista de la física se encuentran:

- **Precisión en las trayectorias atómicas** - Generar trayectorias que coincidan con las trayectorias “reales” seguidas por los átomos no sólo es imposible, sino innecesario. Es imposible seguirlas exactamente debido a las siguientes aproximaciones:
 - En los algoritmos de integración (sección 6.2).
 - En el potencial de interacción (sección 7).
 - En el almacenamiento y control de variables (sección D.2.1).

No tiene sentido esperar seguir las trayectorias en forma exacta porque:

- De cualquier manera hemos incluido errores en las condiciones iniciales (sección 5.1).
 - Las ecuaciones tienen comportamientos dinámicos caóticos que tarde o temprano se harán presentes (sección 6.2).
 - Los resultados esperados dependen de comportamientos generales del sistema, y no de la reproducción exacta de las trayectorias (sección 6.3).
- **Reproducibilidad de los resultados** - Un aspecto básico de lo que queremos esperar de cualquier teoría o experimento científico (sección 4.2.1). La reproducibilidad está íntimamente relacionada con la fase de equilibrio del sistema, dado que una vez que se ha alcanzado el equilibrio el sistema habrá perdido toda memoria de las condiciones iniciales. Los problemas aparecen cuando trabajamos con propiedades de lenta convergencia o que tiene estados metaestables en los que pueda quedar atrapado el sistema. Un ejemplo se muestra en la figura 8-1 donde se observa la evolución de la energía cinética para diferentes configuraciones iniciales.
 - **Medición de muchas propiedades con alta precisión**- Conforme más propiedades midamos de un sistema, podemos considerar mejor a la simulación. Sin embargo, la clave es medir algunas propiedades esenciales, dejando el resto como extra. La precisión en la medición de estas propiedades es también un factor importante. Para poder simular más propiedades con mejor precisión es mejor tener simulaciones que (sección 4.6):

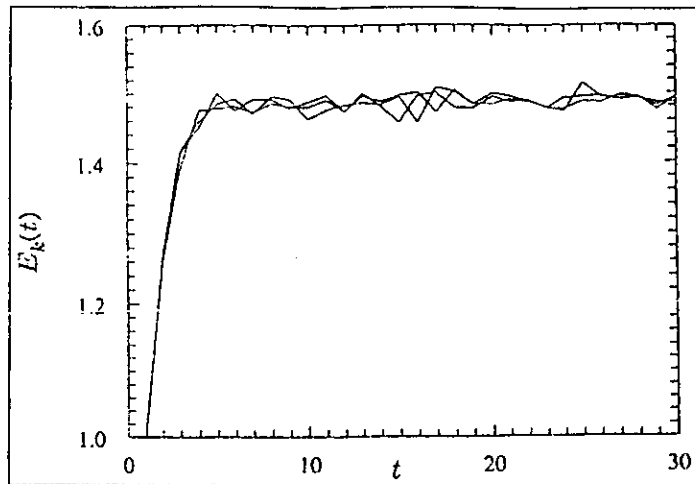


Figura 8-1 Convergencia de la energía cinética promedio desde diferentes estados iniciales. [Rap95]

- Pueda simular tiempos largos - Para poder equilibrar sistemas con lenta convergencia y estudiar la evolución y distribución temporal con mayor espectro y precisión (sección 6.2)..
- Pueda simular sistemas grandes - Para que las aproximaciones en tamaño sean más pequeñas, se puedan tomar en cuenta efectos de grupo mayores y la convergencia a las propiedades de bulto sean mejores (disminuyendo o eliminando los efectos de frontera) (sección 6.3).
- Use potenciales realistas - Para incluir la mayor cantidad de efectos, incluyendo de ser posible efectos cuánticos (sección 7).
- Involucre propiedades relativistas y cuánticas, no sólo de mecánica clásica.

Para que una simulación que nos entrega muchos resultados sea mejor es preferible que requiera de muy pocos datos para llevarse a cabo; no siempre es buena una simulación que nos permita medir 20 variables si es necesario introducir 40 valores.

8.1.2 Bondad como programa

Usualmente la palabra clave al referirnos a la calidad de un programa de simulación es *eficiencia*. Sin embargo, puede referirse a muchas cosas: la velocidad de ejecución, la cantidad de recursos que utiliza, ambas, u otras. Con recursos me refiero a memoria, espacio en disco, papel o cualquier cosa que necesite y se agote.

Que un programa sea eficiente o no es un juicio subjetivo, que depende de muchos factores y puede cambiar de una situación a otra. Por ejemplo, consideremos un programa que utiliza 128K de memoria, requiere 2MB de espacio en disco y tiene un tiempo de ejecución promedio de 7 horas. Si el programa tiene que ordenar una lista con 100 direcciones de correo no es muy eficiente. Sin embargo, si el programa esta ordenando la lista de telefonos de la Ciudad de México seguramente es bastante eficiente.

Es común que optimizar un aspecto de un programa tenga el efecto de degradar otro. Por ejemplo, hacer que un programa se ejecute más rápido usualmente significa hacerlo más grande (desarrollando funciones en lugar de escribirlas por separado); escribir programas más pequeños separandolos en partes los hace más lento. Hacer un uso más eficiente del espacio de disco compactando la información hace más lento el acceso. Este tipo de intercambios entre tipos de eficiencia pueden ser algo frustrantes, especialmente al comienzo cuando no es claro que uno pueda afectar otro.

Afortunadamente existen algunas prácticas de programación más eficientes que otras, así como algunas técnicas que hacen a los programas más pequeños y rápidos. Normalmente estan intimamente relacionadas con el lenguaje de programación que se utilice. Dado que yo programo en C, doy algunos ejemplos de técnicas asociadas al lenguaje C (no necesariamente intransferibles a otros lenguajes) (sección A.2):

1. Usar los operadores de incremento y decremento.

2. Usar variables de registro.
3. Usar apuntadores en lugar de índices en arreglos.
4. Uso de funciones.
5. Usar funciones tan específicas como sea posible (por ejemplo, funciones para enteros al trabajar con enteros).
6. Aprovechar la forma en que la máquina almacena la información para utilizarla en forma óptima [All87, Apéndice A].

Para una descripción detallada sugiero revisar un libro especializado de C, por ejemplo [Ker91],[Str94] y [Sch95].

La eficiencia no lo es todo en un programa, hay ocasiones en que vale la pena sacrificarla, por ejemplo, no vale la pena dedicar meses a optimizar un programa para que necesite uno o dos minutos menos en ejecutar. Algunos valores importantes que puede tener nuestro programa pueden ser:

- Claridad y simplicidad en el código - Hay que hacer notar que claridad y simplicidad en el código no son lo mismo, en ocasiones es preferible usar variables largas y muchas subrutinas para tener claridad, sacrificando la simplicidad. Simplicidad por otro lado sería lo contrario, usar código sin subrutinas que nos ahorra trabajo a costa usualmente de claridad.

En particular el lenguaje C permite una enorme complejidad en el código, siempre es posible escribir lo mismo de una manera diferente, para ganar una millonésima de segundo evitando una llamada a memoria por ejemplo. Si analizamos el código de un programador experto, es normal que nos tome unas horas entender unas cuantas líneas, descifrando la forma en que funciona y la eficiencia que logra. Ciertamente puede ser mucho más eficiente que el código que podamos escribir programadores principiantes, sin embargo no siempre es recomendable seguir el ejemplo. Es indispensable escribir un programa que pueda ser leído y entendido con relativa facilidad. va no digamos por alguien más, sino por nosotros mismos. Puede ser muy frustrante revisar código que hayamos escrito hace muy poco y no entender cómo lo hicimos. Una solución es llenar el programa de comentarios explicativos, usar funciones y variables de nombres "evidentes en si mismos" y mantener un cierto orden. Si uno no sigue estos pasos (y algunos otros que aún no descubro) lo más probable es que, tarde o temprano, se arrepienta.

- Versatilidad - Escribir código que pueda adaptarse a diferentes circunstancias, aunque en el peor de los casos requiera de pequeñas modificaciones.

Es común que en la búsqueda de eficiencia programemos código demasiado específico, y que no podamos utilizar para ninguna otra cosa. No nos interesa comprar una cámara que sólo pueda tomar fotografías personales.

Tampoco es sensato comprar un solo zapato para ambos pies. Hay circunstancias específicas a cada problema que deben tomarse en cuenta y que pueden determinar definitivamente el programa.

- Compatibilidad - Poder aprovechar lo que otros han hecho, por ejemplo, en un programa de simulación poder aprovechar configuraciones de diferentes formatos (sección B) puede extender notablemente el espectro de problemas tratables (y la versatilidad del programa). Crear archivos con los datos de salida que puedan ser aprovechados por diversos programas puede extender nuestras capacidades de analizar resultados.

8.2 La parte más lenta

Si los potenciales de interacción (sección 7) contienen la información física del problema, comenzamos a sospechar que el cálculo de las energías y fuerzas es una de las principales partes de un programa de simulación por dinámica molecular.

Las diferentes partes de la simulación las podemos clasificar en importancia según diferentes valores; quizás los más importantes son

- El tiempo de cómputo requerido por cada una de ellas.
- La dependencia de la evolución y resultados finales a las modificaciones en cada una.

Según ambos valores, los potenciales y el cálculo de sus efectos son definitivamente la parte clave de una simulación. En una simulación simple puede que recurramos al cálculo de interacciones algunos cientos de miles de veces, que puede llegar a ser cientos de millones y más en simulaciones complejas. Un programa puede pasar hasta un 95% del tiempo en esta etapa.

Si utilizamos un algoritmo poco eficiente para generar la configuración inicial probablemente no tenga grandes efectos en el tiempo total de simulación. inclusive si se tarda 5 veces más que un algoritmo eficiente, esto no representa más que una pequeña parte del tiempo total. Por el contrario, si en el cálculo de interacciones utilizamos un algoritmo ligeramente ineficiente, digamos un 20%, puede tener el efecto de reducir en un 19% la eficiencia total del programa. Por lo tanto, si vale la pena optimizar alguna parte de un programa de dinámica molecular, es la de las interacciones entre partículas la mejor opción.

En esta sección expongo algunos de los procedimientos que tiene como objetivo mejorar la eficiencia de un algoritmo de dinámica molecular. Algunas pueden parecer evidentes, pero las incluyo porque puede haber detalles en su implementación que valga la pena comentar.

8.3 Calculando eficientemente las fuerzas, energías y presiones

“Un idiota con computadora es un idiota más rápido”, Rich Julius [Han99].

Si los potenciales más simples son los de pares, comencemos por ellos como primer caso.

Si consideramos dos átomos i y j , usando la convención de imagen mínima (sección 6.3.7), se puede encontrar fácilmente la separación entre ambos, y de ella la fuerza que el átomo i ejerce sobre el j

$$\mathbf{f}_{ij} = -\nabla_{\mathbf{r}_i} \mathcal{V}(r_{ij}), \quad (8.1)$$

en la dirección del vector entre ambos átomos $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. Es fácil demostrar [Gol94] que

$$\mathbf{f}_{ij} = -\frac{1}{r_{ij}} \left(\frac{d\mathcal{V}(r_{ij})}{dr_{ij}} \right) \mathbf{r}_{ij} = -\frac{w(r_{ij})}{r_{ij}^2} \mathbf{r}_{ij}, \quad (8.2)$$

donde $w(r_{ij})$ es la función del virial de pares

$$w(r_{ij}) = r \frac{d\mathcal{V}(r)}{dr} \quad (8.3)$$

utilizada para el cálculo de la presión (sección 6.6.10).

En la ecuación 8.2 se ve que si $\mathcal{V}(r_{ij})$ es una función par de r_{ij} , el vector de fuerza se puede calcular sin calcular la magnitud de \mathbf{r}_{ij} , puesto que r_{ij}^2 es suficiente. Por ejemplo, para el potencial de Lennard Jones (sección 7.4.3) encontramos

$$\mathbf{f}_{ij} = \frac{24\epsilon}{r_{ij}^2} \left[2 \left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \mathbf{r}_{ij}. \quad (8.4)$$

La forma usual de calcular las energías, las fuerzas y la función del virial es con dobles ciclos sobre todos los pares de átomos (equivalentes a las dobles sumas $\sum_i \sum_{i>j}$ presentes en muchas ecuaciones).

Habiendo calculado la fuerza que i ejerce sobre j , la fuerza que j ejerce sobre i se puede calcular fácilmente según la 3ª Ley de Newton que todos conocemos. El detalle está en recordar algunos puntos elementales que pueden hacer este cálculo más eficiente. Para no detenerme en mostrar los detalles hago notar que pueden observarse en [All87, programa F.17], en el libro de Rapaport [Rap95] y en mi programa.

8.4 Evitando la raíz cuadrada

Cuando las interacciones dependen de una potencia impar de r_{ij} (interacciones ion-ion, dipolo dipolo u otras (sección 7)), no hay manera de calcularlas sin encontrar antes la magnitud de la distancia interatómica. Las raíces cuadradas

pueden ser uno de los cálculos más complicados (y tardados) que debamos pedirle a nuestro programa, así que podemos darnos el lujo de sustituirlos por varios cálculos simples y aún así ganar eficiencia.

Cuando utilizamos una celda de simulación relativamente pequeña, como por ejemplo $(-1, +1)$ ó $(-\frac{1}{2}, +\frac{1}{2})$, y las densidades no son muy grandes, podemos utilizar el algoritmo que sugiere Singer [Sin83], cuya validez es aceptable cuando $r_{ij}^2 \in (0.1, 1.0)$. El algoritmo es el siguiente:

Sea $s = r_{ij}^2$.

La raíz cuadrada se expande como un polinomio de mínimos cuadrados, remplazando

$$r_{ij} = \sqrt{s} \quad (8.5)$$

por

$$r_{ij} = c_0 + c_1 s + c_2 s^2 + c_3 s^3 \quad (8.6)$$

donde los coeficientes son [Pow84a] $c_0 = 0.188030699$, $c_1 = 1.48359853$, $c_2 = -1.0979059$ y $c_3 = 0.430357353$. Se pueden utilizar ciclos para refinar la primera aproximación

$$r_{ij} \rightarrow r_{ij}' = r_{ij} + \frac{\delta s}{2r_{ij}} \quad (8.7)$$

donde $\delta s = s - r_{ij}^2$. Tres pasos para refinar dan una precisión de seis dígitos.

Este algoritmo puede llegar a ser 20 veces más rápido que la función SQRT en una CDC 7600, cuando se programa en forma eficiente [All87, Programa F.18]. Sin embargo, hay diferencias importantes de una máquina a otra y puede, incluso, no haber ninguna ventaja.

8.5 Uso de tablas en lugar de funciones

Al complicarse los potenciales usados, evaluar las funciones de potencial y fuerza puede volverse muy complicado (y lento). Una buena solución es utilizar tablas de valores previamente calculados. Usar tablas tiene también la ventaja de que adapta el programa a la posibilidad de utilizar potenciales que no esten dados en forma analítica, como son la mayoría de los más complicados para algunos sistemas moleculares, por ejemplo, el potencial de Barker-Fisher-Watts [Bar71] que requiere de 11 parámetros ajustables y una exponencial.

La tabla se construye una vez al comienzo, y durante la simulación se aprovecha para ahorrar tiempo. Salvo para funciones muy simples es posible ahorrar tiempo con un buen algoritmo para construir y usar tablas.

Para construir una tabla se eligen los límites y la precisión δr . Después se llena la tabla.

Normalmente se construyen varias tablas, una para cada función, de tal manera que puedan moverse en la memoria más fácilmente y se logren cálculos más veloces.

Durante la simulación se requieren valores de la tabla con mucha más precisión de aquella con la que la generamos, y la manera de obtenerla es interpolando los valores conocidos. Podemos usar diferentes algoritmos de interpolación, aunque como la mayoría de los potenciales son suaves (y planos localmente), por lo general no es necesario nada más complicado que una interpolación lineal simple, sobre todo porque interpolaciones complicadas reducen las ventajas, en tiempo, de usar tablas., lineales o no lineales.

Cuando las tablas se usan no sólo para acelerar el programa, sino para utilizar potenciales no definidos en forma analítica (por su definición o por ser leídos de un archivo), es necesario hacer interpolaciones para generar la tabla con la precisión dada. En este caso vale la pena usar un algoritmo no lineal, porque no nos es importante la velocidad (este proceso se realiza solo al comienzo) y podemos sacrificarla a cambio de precisión.

Un ejemplo de interpolación no lineal es el Método de diferencias adelantadas de Newton-Gregory [Boo72], que muestro en la figura 8-2, donde suponemos que la tabla tiene los valores de la energía $\mathcal{V}(s_{k-1}, s_k, s_{k+1} \text{ y } s_{k+2})$ donde $r = r_{ij}^2$. Según el método el valor de $\mathcal{V}(s)$ cuando $s_k < s < s_{k+1}$ es

$$\mathcal{V}(s) \approx \mathcal{V}_k + \xi \delta \mathcal{V}_k + \frac{1}{2} \xi (\xi - 1) \delta^2 \mathcal{V}_k \quad (8.8)$$

cuando

$$\xi = \frac{s - s_k}{\delta s} \quad (8.9)$$

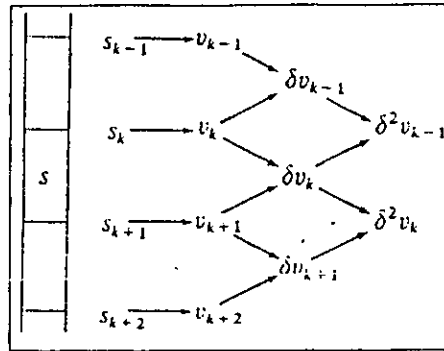


Figura 8-2 Método de Newton-Gregory de interpolación por diferencias adelantadas. [All87]

Además de las energías necesitamos evaluar las fuerzas. Podemos calcularlas construyendo tablas separadas con los valores de la función $w(r_{ij})/r_{ij}^2$, que por medio de la ecuación 8.2 permite calcular las fuerzas. Más simple es utilizar la relación

$$\frac{w(r_{ij}^2)}{r_{ij}^2} = \frac{w(s)}{s} = \frac{2dV}{ds} \quad (8.10)$$

y obtenerla derivando la ecuación de interpolación (por ejemplo 8.8), unos 10,000 puntos por unidad por tabla es común.

Elegir adecuadamente la precisión es fundamental para que usar tablas de resultados correctos. Una manera de elegir δs es con $\delta s = \delta r_{ij}^2 = 0.01 r_m^2$, donde r_m es la posición del mínimo de la función (en este caso el potencial), lo que da 1,000 puntos por unidad por tabla, una cantidad manejable en computadoras modernas (sección A.1.2). En ocasiones no se elige la precisión, sino que se calcula dependiendo de la máxima cantidad de puntos que queremos en la tabla y del intervalo de interés.

Allen describe un método más eficiente en memoria propuesto en [And83], que no requiere de precisión constante y que puede ser importante cuando se usan varios potenciales y no queremos usar tanta memoria en las tablas.

8.5.1 Suavizando datos discretos

Algunos potenciales no los conocemos mediante la función analítica que los describe, sino mediante un conjunto de puntos. Estos puntos son del tipo $(r, V(r))$ y pueden haber sido obtenidos de diferentes formas, experimentalmente, cálculos, graficando funciones analíticas, etc. Cuando conocemos un potencial de esta forma podemos utilizarlo para hacer simulaciones leyendo el archivo con los puntos y almacenándolos en una tabla en la memoria de la computadora. La forma en que aprovechamos estos valores es idéntica a aquella que usamos cuando tenemos potenciales analíticos que queremos utilizar de tablas para ahorrar tiempo de ejecución.

Problemas.

Sin embargo, cuando un potencial esta almacenado en un archivo no necesariamente esta en una forma directamente útil en un programa de simulación:

1. **Precisión.** En el archivo la separación entre los datos puede no tener la precisión que buscamos en la tabla en el programa. Incluso pueden tener separaciones no homogéneas, lo que no es común en las tablas de memoria.
2. **Escalones.** Puede haber sido construido sin tomar en cuenta que las derivadas sean continuas. Este último caso es común cuando el potencial ha sido digitalizado de una imagen. La mayoría de los programas que utilizan extraen una serie de puntos de una imagen no se preocupan por las derivadas de los puntos, sino que digitalizan los puntos que les parecen más cercanos al dibujo. De esta forma se obtienen archivos cuyos puntos forman una serie de escalones. La figura 8-3 muestra estos escalones en un potencial obtenido por digitalización.

Soluciones.

Los archivos leídos deben ser tratados para solucionar estos dos defectos (precisión y derivadas o escalones) antes de ser utilizados. Las soluciones que utilizamos son las siguientes:

1. **Precisión.** Los puntos leídos del archivo son interpolados para encontrar todos los faltantes en caso de que la precisión requerida sea mayor que la almacenada. Si la almacenada es mayor, los puntos sobrantes son ignorados. Cuando las separaciones en el archivo no son homogéneas se utilizan los métodos de interpolación o se ignoran los valores.

Las interpolaciones se pueden hacer con dos precisiones diferentes: lineal o cúbica por partes. En la interpolación lineal se hace una aproximación con una recta que pasa por dos puntos adyacentes al que queremos encontrar. Tiene la ventaja de ser muy simple pero si la precisión no es muy grande introduce errores que pueden tener efectos notables más tarde.

En una interpolación cúbica por partes (*cubic spline* en inglés) se construyen polinomios cúbicos de interpolación para cada segmento, pero además se aplica un algoritmo que garantiza que la primera derivada de los polinomios sean continua en los puntos de intersección. De esta forma el algoritmo permite construir tablas de potenciales en los que las fuerzas también son continuas. Este es el método que preferimos porque garantiza mayor precisión, tanto en la interpolación de los datos como en la construcción de la tabla de fuerzas [Gou96] [Pre86]. Utilizamos el algoritmo descrito en [Gou96] adaptandolo del Basic al C++.

2. **Escalones.** Los escalones los tratamos aplicando un algoritmo recursivo de suavizado. Existen diferentes algoritmos conocidos para suavizar datos, cada uno optimizado para algún tipo de problema. En nuestro caso necesitamos un algoritmo que elimine saltos sin modificar drásticamente los puntos críticos.

El algoritmo que elegimos fué uno de promedio móvil. En este algoritmo se toman grupos de m puntos alrededor de un valor x_i y se promedian de forma pesada, asignando a x_i el valor promedio, aplicandose para toda i entre 0 y el número de puntos en la tabla. Se pueden utilizar diferentes valores para m , pero por lo general se mantiene por debajo de 10. Los coeficientes que determinan el peso de cada punto son simétricos (idénticos para puntos posteriores y anteriores) y pueden generarse de diferentes formas. La que nosotros usamos disminuye cuadráticamente el peso de un punto conforme se aleja del punto de interés. Para los casos $m = 3$ y $m = 5$ las fórmulas usadas son:

$$y_i = \frac{1}{10}(y_{i-2} + 2y_{i-1} + 4y_i + 2y_{i+1} + y_{i+2}) \quad (8.11)$$

$$y_i = \frac{1}{4}(y_{i-1} + 2y_i + y_{i+1}) \quad (8.12)$$

con fórmula general ($\sigma = \frac{1}{2}(m - 1)$ si m es impar y $\sigma = \frac{1}{2}m$ si m es par):

$$y_i = \frac{1}{2^\sigma + \sum_{j=1}^{\sigma} 2^{\sigma-j+1}} (2^\sigma y_i + \sum_{j=1}^{\sigma} 2^{\sigma-j} (y_{i-j} + y_{i+j})). \quad (8.13)$$

El algoritmo implica un recorrido sobre todos los puntos, para luego repetirse, logrando lo que llamamos una *iteración* de suavizado. La figura 8-3 muestra el potencial de Bristowe para Níquel según se obtiene de un archivo creado por digitalización. Son evidentes las discontinuidades en las derivadas. La misma figura muestra el efecto de aplicar un suavizado por promedio móvil como el sugerido. Conforme el número de iteraciones aumenta se puede ver que las derivadas serán más continuas. Sin embargo la suavidad se ha conseguido a costa de un cambio en los parámetros del potencial, como son la distancia de equilibrio y la energía de equilibrio. Se puede ver que cuando entre 100 y 1000 iteraciones se obtienen curvas suaves con variaciones en los parámetros porcentualmente pequeñas.

El efecto de usar diferente número de puntos m en los promedios lo podemos ver en la figura 8-4, en la que se comparan los métodos de 3 y 5 puntos para 10.000 y 36.000 iteraciones. Las diferencias se acentúan conforme se aplican más iteraciones, por ello es que mostramos un ejemplo con tantas iteraciones. Se concluye de la figura que al utilizar más puntos los parámetros cambian más notablemente. Por otro lado, aumentar el número de puntos favorece mejorar la suavidad con menos iteraciones. De esta forma encontramos que para cada caso hay que encontrar un balance entre el número de iteraciones y de puntos para minimizar los cambios en los parámetros mejorando la suavidad.

Para una discusión de diferentes métodos de suavizado para problemas de digitalización recomiendo [Pre86].

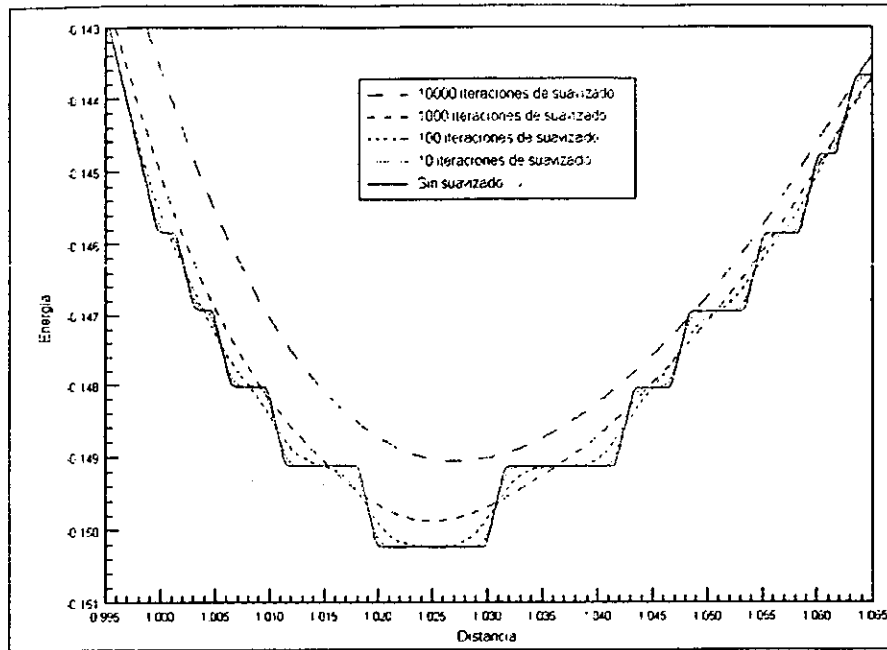


Figura 8-3 Efecto de aplicar diferente número de iteraciones de suavizado al potencial de Bristowe para Níquel generado por digitalización de una imagen. En este caso se usa $m = 5$.

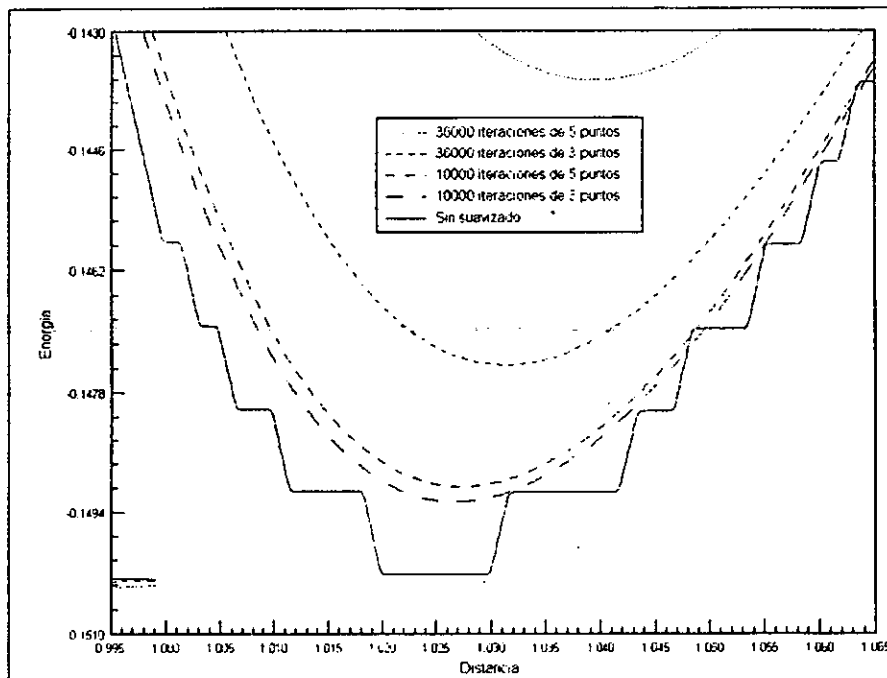


Figura 8-4 Efecto de aplicar diferente número de iteraciones de suavizado al potencial de Bristowe para Níquel generado por digitalización de una imagen. En este caso se usa $m = 5$. Efecto de utilizar diferente número de puntos para aplicar un suavizado por promedio móvil pesado. El potencial original es el mismo de la figura 8-3.

8.5.2 Derivadas numéricas

En ocasiones es necesario calcular numéricamente la derivada de alguna función, por ejemplo porque solamente conocemos la función en puntos aislados. Este caso se presenta cuando utilizamos fuerzas y potenciales en tablas. Cuando los potenciales los obtenemos de un archivo no tenemos acceso a las derivadas para calcular la fuerza, y esta debemos encontrarla nosotros con alguno de los siguientes métodos de derivación numérica [Koo90]. Otra circunstancia en la que es común que surga la necesidad de calcular derivadas en forma numérica es cuando las utilizamos para evaluar propiedades del sistema, donde por supuesto, no conocemos la función analítica.

Para encontrar primeras derivadas [Koo90]:

Hacia adelante de 2 puntos:

$$f' = \frac{f_1 - f_0}{h} \quad (8.14)$$

Hacia atrás de 2 puntos:

$$f' = \frac{f_0 - f_{-1}}{h} \quad (8.15)$$

Simétrico de 3 puntos:

$$f' = \frac{f_1 - f_{-1}}{2h} \quad (8.16)$$

Simétrico de 5 puntos:

$$f' = \frac{1}{12h} [f_{-2} - 8f_{-1} + 8f_1 - f_2] \quad (8.17)$$

Segundas derivadas:

Método de 3 puntos:

$$f'' = \frac{f_1 - f_0 + f_{-1}}{h^2} \quad (8.18)$$

Método de 5 puntos [Dev94]:

$$f'' = \frac{1}{12h^2} (-f_{-2} + 16f_{-1} - 30f + 16f_1 - f_2) \quad (8.19)$$

Terceras derivadas:

$$f''' = \frac{1}{2h^3} (-f_{-2} + 2f_{-1} - 2f_1 + f_2) \quad (8.20)$$

De los anteriores métodos es común utilizar los simétricos, porque siendo simples por lo general son más precisos que los no simétricos.

Al elegir un método debemos tener en cuenta la acumulación de errores que hemos comentado en otras partes del texto (apéndice D.2), ya que conforme h decrece el error decrece también, pero hasta llegar a un mínimo, a partir del cual comienza a crecer nuevamente. Como ejemplo, la fórmula simétrica de 5 puntos para calcular primeras derivadas es más precisa que la de 3 puntos para valores de h más pequeños que 10^{-5} y mayores que 10^{-3} , pero menos precisa que aquella cuando $h \in (10^{-5}10^{-3})$.

Si utilizamos potenciales leídos de un archivo sin suavizar y los derivamos para calcular las fuerzas observamos problemas con la continuidad. Un ejemplo muestro en la figura 8-5, en donde están graficados el potencial de morse y el negativo de su derivada numérica calculada con la ecuación 8.16. Mientras que los puntos del potencial dan lugar a una curva aparentemente suave, los triángulos de la derivada forman una nube que no permite definir claramente los valores adecuados. Si derivamos la fuerza se hacen evidentes los problemas.

En la sección 10 utilizamos el método de 5 puntos para segundas derivadas para encontrar la curvatura de la función de energía de amarre contra radio de Wigner-Seitz. Cuando hicimos esto observamos como utilizar el valor mínimo posible para h no era lo óptimo, y tuvimos que estudiar el efecto de usar h 's mayores. La figura 8-6 muestra la segunda derivada de la función y su dependencia del valor de h . En la figura h es la separación entre puntos (que es constante), mientras que con $2h$, $3h$ y $4h$ designo tamaños mayores. El resultado notable que muestra la figura es que en este caso vale la pena utilizar $2h$ en adelante. Las variaciones en la segunda derivada al calcularlas con h mínima son porcentualmente representativas del valor total. La región denotada en la figura es justo aquella de interés para calcular la curvatura para un potencial de Lennard-Jones.

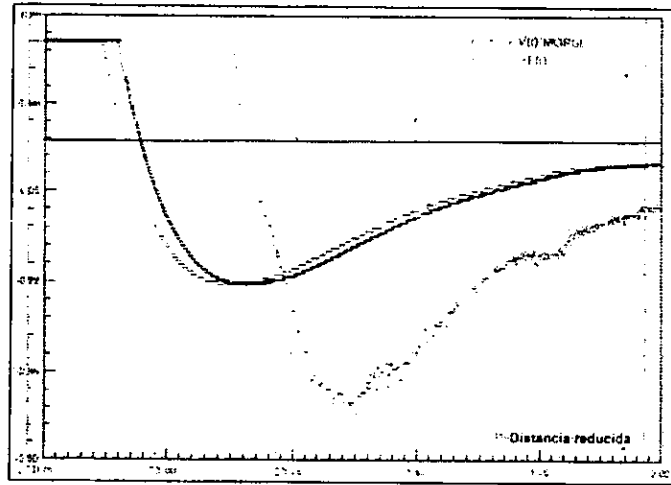


Figura 8-5 Potencial de archivo de Morse (círculos) y el negativo de su derivada (triángulos). Se observan los problemas de calcular derivadas numéricas sin aplicar algoritmos de suavizado. Las discontinuidades en la fuerza harán imposible un cálculo preciso de segundas derivadas.

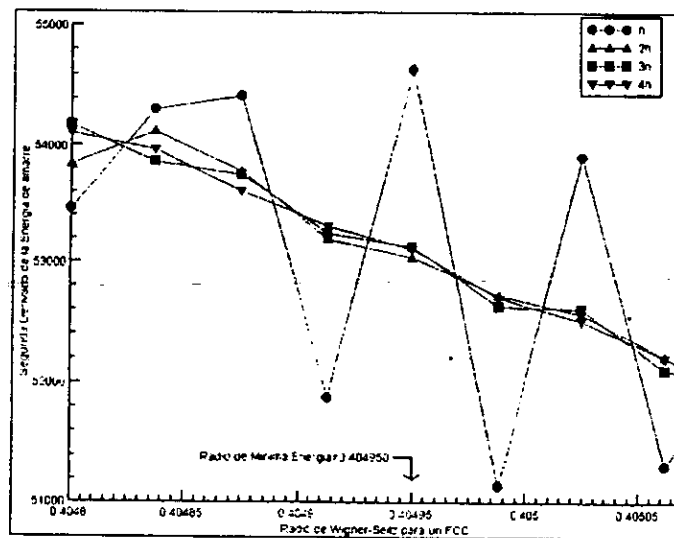


Figura 8-6 Segundas derivadas de la energía de amarre de un cristal FCC según un potencial Lennard-Jones. Note que las variaciones usando h mínima representan una proporción importante. En la figura vemos un detalle de la región correspondiente al mínimo de energía. El método usado es el simétrico de 5 puntos descrito por la ec. 8.19.

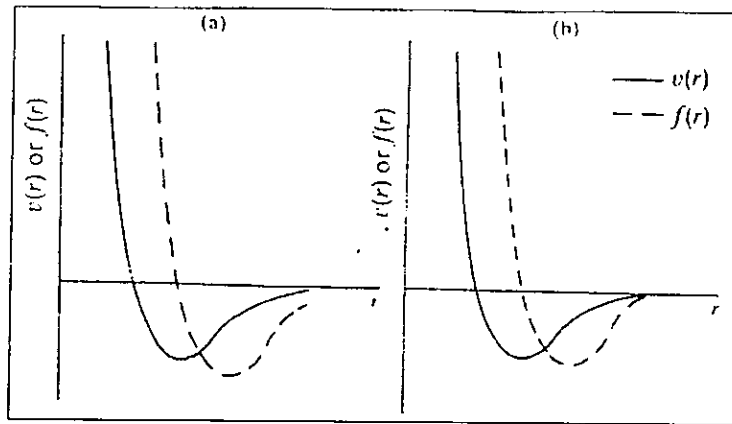


Figura 8-7 Potencial y fuerza de Lennard-Jones a) originales y (b) desplazados. [All87]

8.6 Potenciales desplazados

En la sección del cálculo de la energía potencial (6.6.1) explique que al elegir un radio de corte es importante hacerlo de tal manera que tanto el potencial como sus derivadas convergan suavemente a cero en el punto de corte, de lo contrario habrá una discontinuidad en la energía (de $\mathcal{V}(r_{ij} = r_c)$) cada vez que un átomo entre o salga de la región considerada como de interacción (la esfera de radio r_c).

No siempre es posible elegir radios de corte suficientemente grandes, quizás porque de hacerlo tendríamos que considerar demasiados átomos con el consecuente costo en tiempo; en estos casos una solución posible es desplazar el potencial (y demás funciones derivadas) de tal manera que su convergencia al cero sea suave. El desplazamiento que corrige el problema es $\mathcal{V}_c = \mathcal{V}(r_c)$, transformando el potencial en

$$\mathcal{V}^{\text{desplazado}} = \begin{cases} \mathcal{V}(r_{ij}) - \mathcal{V}_c & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases} \quad (8.21)$$

El término extra en la energía es constante para cualquier interacción, por lo que no afecta las fuerzas, sin embargo, su contribución a la energía varía de paso a paso, porque varía el número de átomos dentro del rango. Es necesario tomar esto en cuenta al calcular la energía total, para poder verificar la continuidad.

Aún existe el problema de la discontinuidad en las fuerzas, que puede dar lugar a inestabilidades en la integración. por lo que algunos autores. [Sto73] [Nic79] [Pow82] [Str78a] usan potenciales de fuerzas desplazadas como este

$$\mathcal{V}^{\text{fuerza desplazada}} = \begin{cases} \mathcal{V}(r_{ij}) - \mathcal{V}_c - \left(\frac{d\mathcal{V}(r_{ij})}{dr_{ij}} \right)_{r_{ij}=r_c} (r_{ij} - r_c) & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases} \quad (8.22)$$

por lo que la discontinuidad no se encuentra ahora en la fuerza, sino en su gradiente. $\left(\frac{d\mathcal{V}(r_{ij})}{dr_{ij}} \right)$

La figura 8-7 muestra el potencial de Lennard Jones, y su fuerza derivada, antes y después de la aplicación del método.

Utilizar potenciales desplazados significa abandonar las ideas físicas que nos llevaron a proponerlo, y que el modelo simulado ya no corresponde con el original, sin embargo, las propiedades termodinámicas de sistemas interactuando con potenciales no desplazados puede obtenerse de simulaciones con potenciales desplazados por medio de un método de perturbaciones [Nic79] [Pow84b].

8.7 Eligiendo el radio de corte

En las secciones 6.3.7 y 6.12.1 se mostraron las virtudes de utilizar un radio de corte r_c en los cálculos de energía. También discutimos métodos para corregir los efectos no deseados de este artificio (secciones 7.8 y 8.6). Para encontrar un

Energía potencial	Radio de corte	Usando la corrección 6.55
-864.78	2.5	-948.72
-920.10	3.5	-946.09

Tabla 8.1 Efectos del radio de corte y corrección en la energía potencial

balance adecuado entre las ventajas y los problemas que encontramos al utilizar los radios de corte y poder decidir un valor adecuado a utilizar voy a mostrar en esta sección más resultados relacionados al problema.

Lo primero es recordar que no existe una fórmula o receta para saber el radio de corte adecuado para un problema. Depende de las necesidades específicas y de las preferencias particulares de cada persona, y de las características del problema.

El efecto de elegir un radio de corte puede ser corregido al utilizar correcciones de largo alcance, como las de la ecuación 6.55, cuyos efectos se muestran en la tabla 8.1.

El problema de utilizar radios de corte es que se desprecian los efectos del potencial más allá de r_c . Sin embargo no siempre introduce un error incluirlo. Por ejemplo, si simulamos un cluster en el que las distancias máximas entre dos átomos son simple menores a 10 unidades de distancia, entonces utilizar un radio de corte de 10 es perfectamente aceptable y no ha de introducir ningún error. Cuando estudiamos sistemas infinitos, como por ejemplo cristales, es imposible no introducir errores de aproximación, porque por la forma en que calculamos las interacciones siempre nos limitamos a cálculos finitos.

Para potenciales de pares la energía total es la suma de las interacciones de pares, así que si minimizamos las aproximaciones en el cálculo de la energía de un átomo se minimizan para el cálculo total.

No es suficiente conocer la distancia a la que un potencial es menor que cierta cantidad que definimos como despreciable para definirla como el radio de corte, porque la energía que estamos despreciando no es ella, sino el producto de esa energía por el número de átomos a esa distancia. Puede haber muchos átomos a una distancia de algunas unidades, aumentando el error.

Si conocemos la distribución de pares en un material y el potencial con el que interactúa, podemos aplicar el potencial a la distribución y calcular la energía total del sistema. En la figura 5-6 de la sección 5.1 mostramos la distribución de pares para una configuración cristalina FCC. Dado que la distribución es discreta es fácil calcular la energía asociada a cada distancia y multiplicarla por el número de átomos a esa distancia para encontrar la contribución de esa distancia a la energía total. La figura 8-8 muestra el resultado de hacer este cálculo para el potencial de Lennard-Jones (y corregir para mostrar las contribuciones en forma porcentual). En la figura se corrigieron las contribuciones de órdenes de vecinos a los que no hay átomos (ver sección 5.1) haciéndolas 10^{-10} para poder calcular el logaritmo. En la figura se puede ver que las contribuciones porcentuales más importantes son las de los primeros ordenes de vecinos, disminuyendo conforme aumenta la distancia.

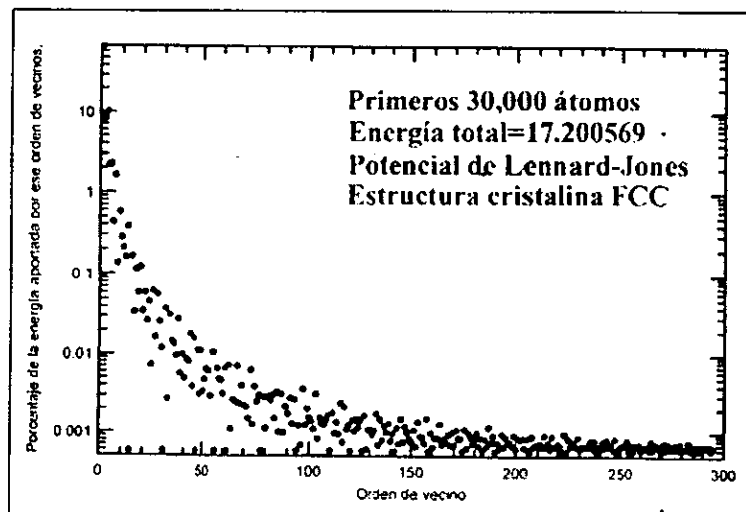


Figura 8-8 Energía potencial como función del orden del vecino.

Con la figura 8-8 podemos elegir la energías que estamos dispuestos a despreciar y encontrar el radio de corte con el que podemos hacerlo. También se puede hacer lo inverso, definiendo r_c y encontrar la energía que estamos despreciando. La figura 8-8 no es de una curva continua, sino una serie de puntos, cada uno resultado de calcular

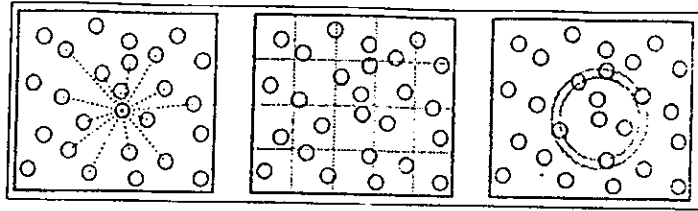


Figura 8-9 Diferentes aproximaciones para calcular interacciones: todos los pares, subdivisión en celdas (de lado mayor o igual que el radio de corte) y listas de vecinos (los círculos concéntricos muestran el radio de corte y el área de donde se toman los vecinos). [Rap95]

la energía de de esa distancia, multiplicarla por el número de vecinos a esa distancia y calcular la contribución porcentual del total. En el eje x no se muestra la distancia sino el orden de vecino, en la sección 5.1 describo la relación interesante entre el orden del vecino y la distancia.

8.8 Listas de vecinos

Al calcular la distancia entre dos átomos elegimos el i y hacemos un ciclo sobre j que recorre todos los demás. Si se encuentran separados por una distancia mayor que r_c , entonces evitamos el cálculo de energías y fuerzas, ahorrando tiempo al no hacer operaciones tardadas.

La cantidad de distancias que debemos examinar según este esquema es proporcional a N^2 (exactamente N^2 si no aprovechamos la 3ª Ley de Newton y $\frac{1}{2}N(N-1)$ si la aprovechamos). Cuando el número de partículas aumente, la cantidad de interacciones crece enormemente, por lo que es conveniente idear alguna forma de reducirla, sobre todo cuando sabemos que para sistemas grandes y potenciales de corto alcance la mayoría de las distancias serán mayores que r_c .

En 1967 Verlet [Ver67] sugiere una técnica para mejorar la velocidad reduciendo el número de distancias calculadas manteniendo una lista de los vecinos más próximos de cada átomo y actualizarla a intervalos. Entre actualizaciones el programa solo calcula las distancias entre un átomo y sus vecinos más cercanos. El número de pares tratados explícitamente se reduce considerablemente. No se ahorra nada de tiempo en el cálculo individual de una interacción, lo que se evita es calcular interacciones poco significativas.

En esta sección describo las técnicas principales para mantener este tipo de listas de vecinos. Rapaport [Rap80] y Erpenbeck y Wood [Erp77] describen una técnica para tratar sistemas de esferas duras, en los que la clave es prever las colisiones moleculares con cuidado; yo no la describo porque la mayoría de las simulaciones hoy en día se hacen con potenciales suaves.

La figura 8-9 muestra las tres técnicas más comunes para reducir el tiempo de cálculo en las interacciones.

8.8.1 Lista de vecinos de Verlet

En el método original de Verlet, la esfera de radio igual al de corte r_c alrededor de un átomo en particular se rodea de un cascarón esférico de ancho Δr , de tal manera que la esfera de radio $r_v = r_c + \Delta r$ contenga a ambas según la figura 8-10. Al comienzo de la simulación se construye una lista de los vecinos de cada átomo que estén a una distancia menor a r_c . Estos se almacenan en un arreglo de tamaño aproximado $4\pi r_v^3 \rho N/6$.

La lista se utiliza en los siguientes pasos de integración para el cálculo de las energías y fuerzas, en donde la manera de identificar los vecinos de cada átomo i varía de un autor a otro. Verlet utiliza un arreglo auxiliar en serie para saber cuales son los vecinos de cada átomo, pero hay muchas opciones con diferentes virtudes y defectos. Para conocer el método de Verlet se puede recurrir a su artículo [Ver67], aunque Allen, Rapaport y Haile [All87] [Rap95] [Hai92] presentan métodos propios equivalentes.

Cada cierto número de pasos la lista de vecinos se actualiza para encontrar los átomos que hayan entrado o salido de la esfera de radio r_c .

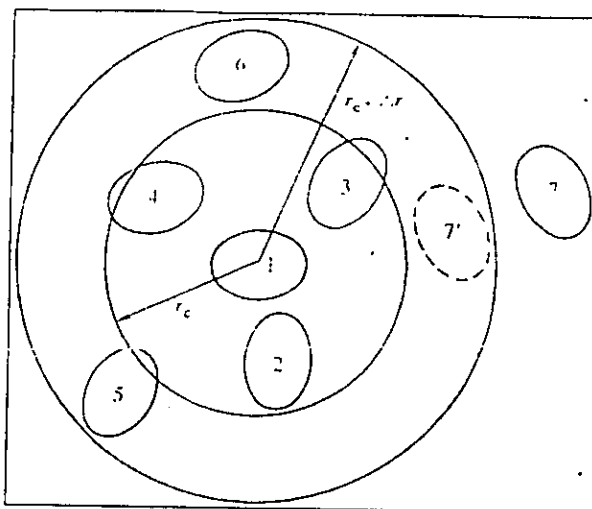


Figura 8-10 La esfera de corte y la cáscara alrededor del átomo 1. Los átomos 2,3,4,5 y 6 están en la lista de vecinos del átomo 1, el 7 no. Sólo los átomos 2,3 y 4 están dentro del radio de interacción al momento de construir la lista. [All87]

El algoritmo se apoya en la suposición de que en muchos de los sistemas simulados existe cierto grado de continuidad a pequeña escala (tanto temporal como espacial), es decir, la vecindad de un átomo se mantiene relativamente constante a lo largo de unos cuantos intervalos de integración. Esto no es fácil de seguir al simular líquidos o algunos gases, pero definitivamente tiene mucho sentido en sólidos. Aún si aceptamos que la vecindad no sea constante, el hecho de incluir en las listas de vecinos átomos que se encuentran en el intervalo de distancias $(r_c, r_c + \Delta r)$, y con los que no se calculan las interacciones, da un mejor grado de confiabilidad a la suposición.

La probabilidad de que la partícula 7 de la figura 8-10 cruce el cascarón de ancho Δr en algunos intervalos dt , puede ser muy baja. No es importante si las moléculas que ya estaban dentro se quedan dentro, pues continuamos usando el radio de corte.

Es común que una lista no necesite actualizarse antes de 10 o 20 pasos. El número de pasos de integración entre actualizaciones de la lista usualmente se fija al comienzo de la simulación, cuando en simulaciones previas se han encontrado valores útiles. También es posible verificar el número de correcciones a las listas de vecinos, de tal manera que durante la simulación se pueda adaptar el tiempo entre actualizaciones en forma dinámica. También es posible tratar de encontrar un momento adecuado para una actualización siguiendo el desplazamiento total de todas las partículas desde la última actualización. Cuando la suma de las magnitudes de los dos mayores desplazamientos sea mayor a Δr , es necesario volver a actualizar [Fin81] [Tho83].

Rapaport [Rap95] sugiere un método más para decidir hacer la actualización: monitorear la máxima velocidad en cada paso hasta que

$$\sum_{\text{pasos}} \left(\max_i |v_i| \right) > \frac{\Delta r}{2\Delta t} \quad (8.23)$$

El criterio es equivalente a examinar los desplazamientos atómicos, es conservador al tomar las contribuciones de todos los átomos, garantizando que no dejaremos de tomar en cuenta ninguna interacción más cercana que r_c .

El criterio anterior es útil en sistemas aislados, pero al aplicar condiciones de periodicidad y hacer pasar un átomo de un lado de la celda original al otro las velocidades no muestran la necesidad de una actualización, la cual es indispensable para evitar cálculos equivocados. No son tan graves los átomos que siguen considerando al átomo trasladado como vecino, pero si aquellos que se encuentran muy cercanos y que no están considerando la interacción. El efecto se observa en fallas catastróficas en la conservación de la energía.

Para evitar estos problemas utilizamos un algoritmo de verificación de necesidad de actualizar listas de vecinos basado en las posiciones. En nuestro algoritmo se almacenan las posiciones de los átomos cada vez que hay una actualización de vecinos. Cuando durante la evolución los átomos se mueven sus nuevas posiciones se comparan con las almacenadas. Si al comparar estas posiciones hay alguna diferencia mayor que cierta función de Δr entonces se hace una actualización.

En resumen, se hace una actualización de vecinos si para algún átomo

$$|\vec{r}_i^t - \vec{r}_{i=t_{\text{últ}}}| \geq \frac{\Delta r}{\gamma} \quad (8.24)$$

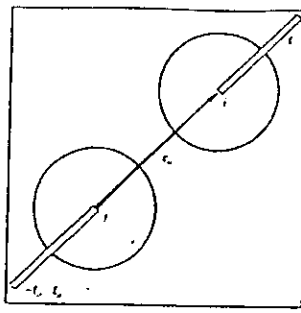


Figura 8-11 El vector de separación y la fuerza entre dos átomos. [All87]

donde \vec{r}_i^j es la posición al tiempo t y $\vec{r}_{i=tc}^j$ la posición al tiempo de actualización. El valor de γ determina que tan sensible se es a hacer actualizaciones, y encontrar el valor adecuado para un sistema requiere de pruebas. El valor estándar de γ es de 2, sin embargo es posible que valga la pena hacer actualizaciones más seguido, pero depende del tamaño del radio de corte.

Utilizar un valor grande para Δr permite no tener que actualizar muy seguido las listas, sin embargo la eficiencia del algoritmo por paso empeora. Utilizar un Δr muy pequeño otorga la máxima eficiencia por paso, pero puede requerir actualizaciones tan seguido que nuevamente la eficiencia total empeore. Thompson [Tho83] estudio el balance adecuado, y sus resultados los resumo en la figura 8-11. Lo que aquí he denominado $r_v = r_c + \Delta r$ es lo que en la figura 8-11 llaman *Radio de la lista* (o *List Radius* en inglés). En un sólido como los que dan lugar a las fronteras de grano estudiadas en esta tesis, los átomos se mantienen relativamente fijos en sus lugares iniciales, y es posible que, aún para radios pequeños y eficientes, no sean necesarias actualizaciones a la lista.

En la sección 8.3 sugiero el uso de la 3ª Ley de Newton para calcular la fuerza que el átomo j ejerce sobre el átomo i , cuando sabemos la que el átomo i ejerce sobre el j . De esta forma, una lista de vecinos del átomo i no necesita incluir más que los átomos j con $j > i$. Cuando el sistema crece, aún cuando las listas se hacen muy grandes el método contribuye más, justificando su implementación.

8.8.2 Subdivisión en celdas

Una manera de reducir el número de interacciones a calcular a un orden N es utilizar el algoritmo de subdivisión en celdas [Sch73] [Hoc74], que consiste en lo siguiente: Dividamos la región simulada (celda original) en pequeñas sub celdas de tal manera que todos los lados de estas sean mayores que r_c . Asignamos ahora los átomos a la celda en que se encuentran. Por el tamaño de las sub celdas, las interacciones sólo son posibles entre átomos de una misma sub celda o sub celdas adyacentes. Por simetría (nuevamente la 3ª Ley de Newton) solamente es necesario considerar la mitad de las sub celdas adyacentes, 14 en un sistema de tres dimensiones y 5 para un sistema de dos dimensiones (incluyendo la sub celda que contiene al átomo en cuestión). Las condiciones de frontera periódicas son fáciles de aplicar en estas circunstancias, aunque hay que tener cuidado al implementarlas. Lo común es suponer que la celda original se ha dividido, al menos, en tres celdas en cada dirección (9 en total).

Es necesario que la región estudiada mida al menos $4r_c$, pero generalmente esto no es difícil.

La forma común (más no indispensable) de programar un cálculo de fuerzas en un sistema dividido como el que creamos involucra una manera de organizar la información conocida como *lista enlazada* [Knu68]. No describo el método, pero una buena explicación puede encontrarse en los libros de Allen [All87] o de Rapaport [Rap95], sin embargo, me anotar un comentario de Rapaport sobre la virtud del método de lista enlazada: Reservar el espacio necesario para almacenar los átomos contenidos en cada una de las sub celdas permitiendo que cualquiera de ellas tenga el máximo empaque desperdicia mucha memoria. Por el contrario, la lista enlazada permite conocer con anticipación la forma en que será requerida la memoria.

La implementación es más larga y compleja que la interacciones todos contra todos, pero es mucho más rápida para cualquier sistema más allá del tamaño mínimo.

En ocasiones se mejora aún más la eficiencia de este método organizando las listas enlazadas de átomos de cada sub celda, por ejemplo de una esquina a la otra, de tal manera que pueda ser más simple saber cuando ya no puede haber vecinos suficientemente cercanos.

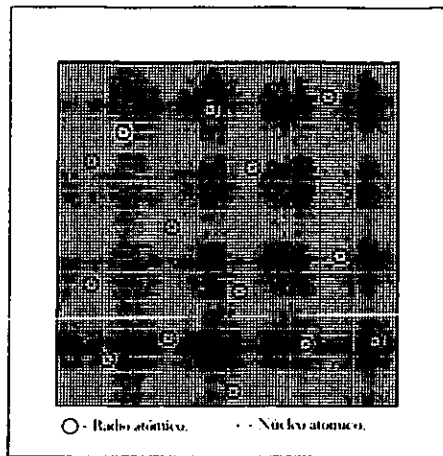


Figura 8-12 En la figura muestro la subdivisión en celdas mínimas. Cada cuadro es una posición posible, los cuadros rellenos representan un núcleo atómico, mientras que los círculos representan la distancia a la que la fuerza de repulsión es muy grande.

8.8.3 Celdas mínimas

Hay un método muy semejante al de subdivisión en celdas y que me gustaría mencionar.

La idea es subdividir la celda original en celdas suficientemente pequeñas como para garantizar que sólo hay un átomo dentro. Para potenciales de esferas duras esto corresponde con el radio de interacción, pero para potenciales suaves puede ser aquel radio para el cual la energía sea ridículamente grande (digamos, altamente improbable). De esta forma no requerimos listas enlazadas, y se calculan las interacciones entre todas las celdas a distancias menores que r_c . (figura 8-12).

Una ventaja de este método es que la lista de celdas a distancia menor que r_c no necesita ser actualizada jamás, la distancia puede estar almacenada junto con ella, y para incluir la interacción solo es necesario saber si la celda esta ocupada o no.

Sería posible hacer una simulación con algún otro método y almacenar las distancias mínimas a las que llegan a encontrarse los átomos, incluso a grandes temperaturas para disminuir esta distancia, y luego utilizar esta distancia mínima como base de la subdivisión.

8.9 Método de muchos intervalos

He descrito ya diferentes métodos para evitar calcular interacciones con átomos relativamente lejos, que logran reducir el tiempo de cálculo notablemente. Vale la pena ahora intentar reducir el tiempo total reduciendo el necesario para calcular interacciones *dentro* de la esfera de corte. Para esto se ha pensado el método de muchos intervalos.

El método de muchos intervalos es particularmente útil para tratar con fuerzas de medio alcance, para las que las correcciones de largo alcance no son tan necesarias y los radios de corte relativamente grandes hacen poco útiles las listas de vecinos.

La figura 8-13 muestra la vecindad de un átomo dividida en dos partes: la primera, delimitada por una distancia r_p y la segunda entre las distancias r_p y r_s . Separamos la fuerza en dos componentes, dependiendo de si provienen de interacciones con átomos de la primera o la segunda region (llamados *vecinos cercanos* y *vecinos lejanos*). Las interacciones más importantes y variables son aquellas con los vecinos cercanos, mientras que aquellas con los vecinos lejanos, aunque notables son más constantes. La intención es tomar en cuenta esta diferencia y calcular mas seguido las primeras y menos las segundas.

Una opción es calcular las interacciones con los vecinos cercanos cada paso de tiempo, y aquellas con vecinos lejanos cada cierto número de pasos. Las interacciones entre los cercanos se pueden calcular de las formas usuales, pero podemos mejorar los cálculos con los lejanos de esta forma: calculando no sólo la fuerza, sino algunas de sus

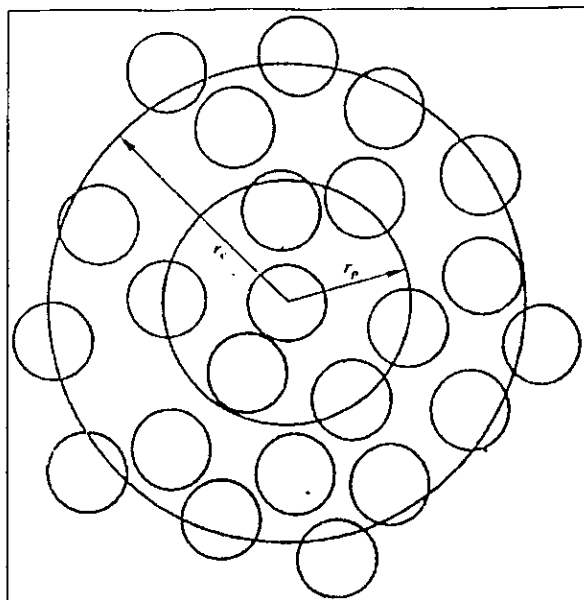


Figura 8-13 Los vecinos cercanos de un átomo, divididos en primarios y secundarios. [All87]

derivadas y funciones de las posiciones de los vecinos lejanos en un instante, y aprovechar esta información para extrapolar los valores de la fuerza en los intervalos que siguen, hasta llegar al siguiente intervalo de cálculo.

Debido a que el método utiliza derivadas de varios ordenes es particularmente adaptable a algoritmos de integración de tipo predictor-corrector (sección 6.2.9).

Según Streett [Str78a] las ganancias de aplicar un método como este pueden ser de un factor de 7 a 10 comparado con un programa de dinámica molecular común, y de 3 a 5 comparado con un programa que utiliza una lista de vecinos de tipo Verlet.

Ha sido aplicado a problemas de fluidos y cadenas moleculares ([Str78b] y [Swi84] respectivamente), y extendido a tres regiones de interacción. Aún hay dificultades en su aplicación, porque sólo ha logrado hacerse con éxito para intervalos dt relativamente pequeños ($5 \times 10^{-15}s$ comparado con $10^{-14}s$ de la dinámica molecular usual [Ver67]).

8.10 Usando varios procesadores

Hasta aquí he descrito solamente métodos para aumentar la eficiencia de un programa de simulación. Podemos rompernos la cabeza tratando de aumentar la eficiencia utilizando procedimientos más complicados, pero es necesaria una estructura diferente tanto en la computadora como en el programa para llevar la eficiencia más allá.

La solución son las nuevas computadoras con varios procesadores. Para programar de forma efectiva para ellas es necesario entender cómo es que las máquinas procesan la información, aunque por lo general es suficiente un tratamiento superficial.

No voy a extenderme en el asunto porque no es el tema del trabajo. Es común referirse a la programación para varios procesadores como paralela.

Para una introducción simple y agradable del tema sugiero el capítulo 14 del libro de Rapaport [?].

8.11 Usando varias computadoras

Es común desarrollar programas con la intención de ejecutarlos en diferentes computadoras, o lo que es aún más complicado, utilizar computadoras diferentes para desarrollarlos. Las diferentes computadoras pueden tener distintos

procesadores, sistemas operativos o ambos. Transportar un programa de una máquina a otra se llama *portar*, y puede ser muy fácil o muy difícil, dependiendo de como se escribió el programa originalmente. Un programa que puede ser portado fácilmente se llama *portable*. Cuando un programa no es fácilmente portable, usualmente es porque es *dependiente de la máquina*, es decir, contiene fragmentos que sólo trabajan con ese sistema operativo o procesador. El lenguaje C permite escribir código portable, pero para conseguirlo es necesario tener cuidado y atención a los detalles.

9. DISEÑO Y FUNCIONES DEL PROGRAMA

9.1 Que esperamos del programa

“Caminar sobre el agua y desarrollar software desde especificaciones es fácil si estan congeladas”.
Edward V Berard [Han99].

Hasta ahora he expluesto diferentes posibilidades y opciones en dinámica molecular. Ya sea en algoritmos, cálculos, valores de dt , y más. La sección anterior describe las características que hacen de un programa un *buen simulador*. Si algo queda claro de todo esto es que ningún programa es el mejor para todo, y que un programa puede ser mejor conforme se haga mas específico al problema de interés. A continuación hago una recopilación de las características que decidimos incorporar a nuestro banco de trabajo. Dado que esta tesis pretendia investigar no sólo las fronteras de grano generadas según el modelo GCSN sino diferentes aspectos de la dinámica molecular que pueden ser definitivos en los resultados. hicimos programas específicos para diferentes problemas. De ahora en adelante me referire a estos programas específicos como *programas menores*, porque son más pequeños, sencillos y sólo tienen como objetivo complementar al *programa principal*.

Todo lo que se dice en esta sección se refiere al programa principal de simulación de dinámica molecular para fronteras de grano. En las secciones correspondientes se discute lo referente a los programas menores.

9.1.1 Lo que si debe hacer

Un programa de dinámica molecular útil al departamento de materia condensada del IFUNAM para simular fronteras generadas según el modelo GCSN debe tener las siguientes propiedades (sin algún orden particular):

- Estar programado en C ó C++ ANSI para poder ser portado a las diferentes máquinas del departamento y mantenido por los integrantes departamento acostumbrados a programar en C (sección 8.11).
- Configuraciones cristalinas FCC o leídas de archivos PDB (sección 5.1.1).
- Velocidades iniciales según una distribución de Maxwell con cierta temperatura (sección 5.1.2).
- Usar los diferentes tipos de condiciones de frontera conocidos. explorando sus efectos (sección 6.3).
- Usar diferentes potenciales, de pares y de átomo embebido, tanto analíticos como de tablas. Calculando las fuerzas y energías durante la ejecución o usando tablas (secciones 7 y 8.5).
- Listas de vecinos, cuya frecuencia de actualización pueda ser calculada en forma dinámica (sección 8.8).
- Un máximo de 20,000 átomos.
- Algoritmo de Verlet velocidad con dt variable (sección 6.2).
- Diferentes algoritmos para llegar a un equilibrio y minimizar energía (sección 6.11).
- Posibilidad de agregar en el futuro más funciones (sección 9.4).
- Facil de mantener y posibilidad de acoplarlo al programa de Monte Carlo del departamento (secciones 9.4 y 4.7).
- Ejecutar simulaciones en algunas horas como máximo (sección 8.2).
- Salida de cálculos en formatos legibles por fgraph (sección 9.9).

- Salida de películas en formato ReView (sección 9.9).
- Un archivo de control completo y fácil de modificar (sección 9.8).
- La posibilidad de agregar más tarde simulaciones de diferentes ensambles (sección 6.14).

9.1.2 Lo que no debe hacer

Así como hay algunas cosas que el programa debe tener, hay algunas que no debe tener:

- Mostrar en tiempo real el sistema en forma gráfica. Por requerir mucho tiempo.
- Usar diferentes algoritmos de integración. Porque Verlet velocidad ha mostrado ser adecuado para el problema. (Finalmente la programación en objetos permite usar diferentes algoritmos, pero no lo hicimos) (sección 6.2).
- Cálculo en tiempo real de propiedades complejas, por ejemplo patrones de rayos x. De necesitarlos se calcularán posteriormente en base a configuraciones grabadas (sección 6.10).

9.2 ¿Que maquina vamos a usar?

Determinadas las necesidades del programa podemos buscar las máquinas mas adecuadas a un programa de este tipo.

Las máquinas del Departamento de materia condensada del IFUNAM son de dos tipos: estaciones de trabajo Silicon Graphics Indy e Indigo 2, y computadoras personales con sistemas operativos varios (Windows 3.11, 95 y NT 4.0). Han sido utilizadas para hacer simulaciones de Monte Carlo de sistemas similares a los que trataremos en esta tesis, mostrando ser adecuadas, tanto en velocidad como en capacidades de almacenamiento [Fer98]. En la sección A.1 discuto diferentes consideraciones para afirmar que una máquina es adecuada.

La UNAM cuenta con varias supercomputadoras, disponibles para simulaciones de dinámica molecular particularmente exigentes. Las principales ventajas son su gran capacidad de procesar en paralelo. Debido a los requerimientos especiales de este tipo de programación en estas máquinas (sección 8.10), decidimos que no valía la pena la inversión de tiempo y esfuerzo, sobre todo porque las máquinas locales al departamento demostraron ser suficientes.

La mayor parte de la programación fué hecha en computadoras personales, con procesadores Intel Pentium II de al menos 266 mhz y 64 MB RAM, con sistema operativo Windows (95, 98 y NT 4.0), usando Visual C++ 5.0 y 6.0. Utilizamos las máquinas Silicon Graphics para ejecutar el programa y estudiar resultados. Nos apoyamos en muchos programas, entre ellos los de la sección A.3. Para una discusión de los efectos del equipo usado favor de ver la sección A.1.

9.3 ¿En que lenguaje?

"C hace fácil dispararse uno mismo en el pie. C++ lo hace más difícil, pero si ocurre vuela la pierna completa". Bjarne Stroustrup creador de C++ [Han99].

En la sección A.2 hay una discusión detallada de las diferentes virtudes y defectos de los lenguajes de programación existentes relacionados con la simulación. Tomando en cuenta estas consideraciones elegimos ANSI C y C++ para este programa, por diferentes razones:

1. Permite aprovechar al máximo la máquina [Ker91].
2. El personal del departamento tiene amplia experiencia en este lenguaje. El programa existente de Monte Carlo está escrito en C++, lo que facilita su integración y el desarrollo de un programa híbrido [Fer98].

3. Los compiladores de C y C++ accesibles en Windows y en Silicon Graphics son más poderosos y cómodos que los accesibles en fortran [Buz94].
4. La interface para mejorar la comodidad puede ser programada con máxima facilidad en Visual Basic, por lo que las dificultades de interface en C no se encuentran [Mic98].
5. La posibilidad que brinda C++ de programar con orientación a objetos y el poder de sus apuntadores [Dev93], [Dev97], [Ceb98] y [Sch95].

Habrán quienes estén en desacuerdo con nosotros por elegir este lenguaje y no el fortran, que en su versión 90 tiene muchas de las virtudes elegidas, y tienen razón porque al abandonar fortran renunciamos a otras ventajas, como por ejemplo la posibilidad de usar las rutinas y librerías existentes en fortran, sin embargo la experiencia ganada al escribirlas en C específicamente para el departamento resulta clave al extraer conclusiones de los resultados.

9.4 Objetos

“Si yo tuviera que vender mi gato (al menos a un informático) no diría que es amable y autosuficiente y que se alimenta de ratones: más bien argüiría que es orientado-a-objetos”, Roger King[Dev97].

Sabemos de las enormes ventajas de la *programación orientada a objetos*, sin embargo, queda pendiente la pregunta ¿que es un lenguaje de programación orientado a objetos?

Algunas personas suponen, de forma bastante lógica, que los lenguajes de programación orientada a objetos se utilizan para crear y gestionar menús, iconos, y los restantes elementos de la interfaz con el usuario que se encuentran en los programas basados en ventanas (Windows por ejemplo). Los programadores en Windows, sin embargo, han creado y gestionado objetos de pantalla durante años utilizando lo que llaman API (Application Programming Interface) de Windows, que no es una herramienta orientada a objetos [And97].

Mucha gente define también un lenguaje orientado a objetos como un lenguaje que hace posible reutilizar código. Esta definición se encuentra un poco más próxima a la verdad, pero no es precisa. Un buen programador en C, por ejemplo, puede escribir funciones en C que sean reutilizables de una aplicación a la siguiente, y un mal programador en C++ puede escribir código que no sea reutilizable en absoluto. Como lenguaje orientado a objetos, C++ ofrece un cierto número de mecanismos de programación que hacen sencillo escribir código reutilizable, pero estos mecanismos no son sus características definitivas.

La forma más útil de comprender C++ como un lenguaje orientado a objetos es comprender cómo C++ saca partido de las siguientes características:

- **Herencia.** En un lenguaje no orientado a objetos, como C, no se puede crear una estructura de datos que herede características de otra estructura de datos. Cada vez que se desea crear una estructura de datos se tiene que partir de cero. En C++ se puede crear una estructura de datos que herede características de otra estructura de datos y que aumente esas características con algunas nuevas propias. Además, usted puede crear funciones (denominadas funciones miembro) que *pertenezcan* a estructuras de datos de C++ y puede crear otras estructuras de datos que hereden esas funciones y utilicen otras funciones propias. En C++ las estructuras que heredan datos y funciones de otras estructuras se encuentran organizadas en *jerarquías* de herencia. Utilizando estas jerarquías, no sólo se puede escribir código que sea reutilizable, también puede escribir estructuras de datos (denominadas *clases*) que contienen variables y funciones que también son reutilizables, totalmente o en parte.
- **Encapsulación de datos.** En C++ las funciones miembro pueden acceder a todas las variables miembro de la misma clase. Sin embargo, un objeto puede proteger sus funciones y variables miembro impidiendo que se acceda a ellas o sean modificadas desde otras clases. La habilidad de que un objeto pueda ocultar sus datos a otras partes de un programa se denomina *encapsulación de datos*.
- **Abstracción de datos.** Al diseñar una clase en C++ se pueden ocultar los detalles de cómo están representados y cómo son manejados. Haciendo uso de la abstracción se pueden ignorar los detalles de cómo es implementada una operación, pudiendo concentrarse en las labores a realizar.

9.4.1 ¿Como se crean?

Un objeto es una instancia de una clase. Así como *cada uno de nosotros* es una instancia de la clase *humano*. Una vez especificadas las características de una clase, aún queda crear los objetos [Dev93].

Un programa en C++ define diferentes clases, pero no es sino hasta la ejecución del programa que en la memoria de la máquina se crean los objetos de estas clases. En nuestro programa aprovechamos esta posibilidad, definiendo objetos de todos los tipos que pueden requerirse en una simulación. En tiempo de ejecución se instancian solamente los objetos necesarios para la simulación específica del momento. De esta forma se hace un uso más eficiente de la memoria de la máquina.

9.4.2 Interacciones entre objetos

Una clase comprende las funciones que representan la manera en que un objeto interactúa con otro. Las clases de nuestro programa contienen datos que deben ser usados y modificados por otras clases, sin embargo la encapsulación de datos protege los datos de una clase. Para salvar este problema al construir una clase incluimos información (apuntadores) a las clases cuyas funciones necesita. De esta forma una clase sólo conoce la forma de las clases que necesita utilizar. Programando las clases de esta forma podemos utilizar funciones en forma independiente de la forma en que están implementadas o de la clase específica a la que pertenecen. Por ejemplo: al aplicar condiciones de frontera podemos simplemente llamar a una función *Frontera*. La función específica que será llamada durante la ejecución puede ser la implementada en las clases *Periódica*, *Libre*, *Espejo* o alguna otra.

Después de instanciar los objetos necesitamos entregar a cada uno la información (dirección en memoria por medio de un apuntador) de los demás objetos creados. La ventaja de hacer explícitamente la entrega de información es que permite una absoluta flexibilidad en programación y ejecución. En tiempo de ejecución tenemos la libertad de instanciar los objetos de las clases que necesitemos, inclusive podemos instanciar diferentes objetos de una misma clase (lo que hasta ahora no hemos necesitado hacer).

9.4.3 Ventajas de usar objetos

Tal y como explico en la sección A.2.2 existen diferentes lenguajes comunmente utilizados para programar simulaciones, pero C++ tiene ventajas importantes sobre la mayoría para programar dinámica molecular. En la sección A.2.3 describo algunas de las virtudes de C++ en comparación con C. La posibilidad de programar con orientación a objetos en C++ nos convencio de utilizarlo. Las principales ventajas de este tipo de programación son [Dev93], [Buz94]:

- La posibilidad de encapsular funciones, disminuyendo la interferencia entre diferentes partes del programa, limitando la posibilidad de dañar partes del código que trabajan bien al modificar otras, dejando siempre abierta la posibilidad de extender la capacidad del programa.
- Las funciones virtuales, que permiten programar y ejecutar de forma mucho más eficiente.
- Mucho más facil extender los procesos y dar mantenimiento al programa.
- Código más elegante y fácil de leer.

Los compiladores de C++ son compatibles también con la versión ANSI de C. de tal manera que la mera utilización de C++ no garantiza aprovechar sus ventajas, es necesario programar utilizando sus capacidades de orientación a objetos.

Esta orientación a objetos también significa un costo, tanto a la hora de programar como a la hora de ejecutar un programa. Algunos autores consideran que la pérdida en tiempo real de ejecución de un programa orientado a objetos con respecto a uno estructurado puede ser tan grande como un 5%. En realidad es muy difícil cuantificar esta diferencia de forma general. Es posible observar una diferencia comparando programas "equivalentes", pero dado que por lo general al pasar de una técnica de programación a otra se modifican notablemente los algoritmos, una comparación directa no es evidente [Buz94].

9.4.4 Desventajas de usar objetos

Hay desventajas de utilizar programación orientada a objetos con respecto a la programación estructurada. Estas desventajas pueden presentarse en las cuatro etapas de la vida de un programa: en la planeación y diseño, en la escritura, en la compilación y en la ejecución. Las desventajas en la escritura y compilación, aunque existen, por lo general son despreciables para consideraciones prácticas porque son balanceadas por las ventajas y porque son muy pequeñas.

Por el contrario, determinar los diferentes objetos y las relaciones entre ellos a la hora de diseñar un programa es esencial para el desempeño de este, y puede ser una muy buena parte del tiempo de preparación de un programa. Si estas elecciones se hicieron de forma adecuada por lo general la escritura y el mantenimiento del programa son directas. Hay muchos libros dedicados al análisis de los objetos y a las técnicas para programar con ellos [Dev93],[Dev97] y [Ceb98]. En la programación estructurada es más directo encontrar diferentes subrutinas y funciones en que dividir el programa, y de allí se hacen evidentes los ámbitos en que deben ser declaradas las variables.

Durante la ejecución las desventajas de la programación orientada a objetos son notables. Al apreciar propiedades de los objetos las funciones involucradas son llamadas según una cadena enlazada de apuntadores, y el seguimiento de estas genera una sobrecarga en procesador.

Para el programa de esta tesis decidimos utilizar la orientación a objetos porque consideramos que las virtudes en cuanto a mantenimiento, versatilidad y claridad del código sobrepasan por mucho a los defectos. Sin embargo, quiero comentar que de ninguna manera es recomendable siempre programar de esta forma. Es más recomendable programar en forma estructurada cuando:

- El programa es muy sencillo (tiene pocas subrutinas y funciones separadas) por lo que la encapsulación y el tiempo perdido en las cadenas asociadas a los objetos no se justifica.
- Queremos la máxima eficiencia de ejecución. No podemos negar que el costo en tiempo al usar objetos existe, por lo que la máxima eficiencia no se puede alcanzar de esta forma. De cualquier manera, si lo que buscamos es una máxima eficiencia la programación con subrutinas y funciones tampoco es la mejor opción. Depende de lo desesperados que estemos por mejorar el rendimiento. En nuestro caso observamos que la única opción que permitía un crecimiento y mantenimiento comodo del programa era la orientación a objetos.

Dado que queremos utilizar el programa en diferentes computadoras, con diferentes sistemas operativos decidimos que debía estar programado en C++ estándar, para poder compilarlo y ejecutarlo sin complicaciones tanto en las PC con Windows 95 y NT, como en las Silicon Graphics con sistema Unix. La plataforma en que desarrollamos el programa es Visual C++ 6.0 de Microsoft para Windows 95/NT. Es una lástima no poder aprovechar la gran cantidad de librerías que este paquete incluye, pero vale la pena a cambio de la posibilidad de transportar el programa de sistema.

9.4.5 Objetos en nuestro programa

Jerarquías de clases

la figura 9-1 muestra la jerarquía de clases, aclarando las relaciones de derivación y herencia.

Clases

1. cAjustes - Tiene la función que prepara los parámetros. Los diferentes tipos de simulación se separan según el tipo de ajustes hechos sobre el sistema en evolución, por ello se derivan de esta clase, la que contiene las funciones que modifican al sistema, como escalar velocidades, centrar el cluster o hacer el momento lineal total cero. Esta íntimamente relacionada con cRecolectarDatos quien la provee de la información necesaria para hacer los ajustes.
2. cAzar - Configura las velocidades con dirección y magnitud al azar (luego son escaladas para determinar la temperatura).
3. cBCC - Llena el cluster con átomos en una configuración cúbica centrada en la cara. El cristal se construye celda por celda.

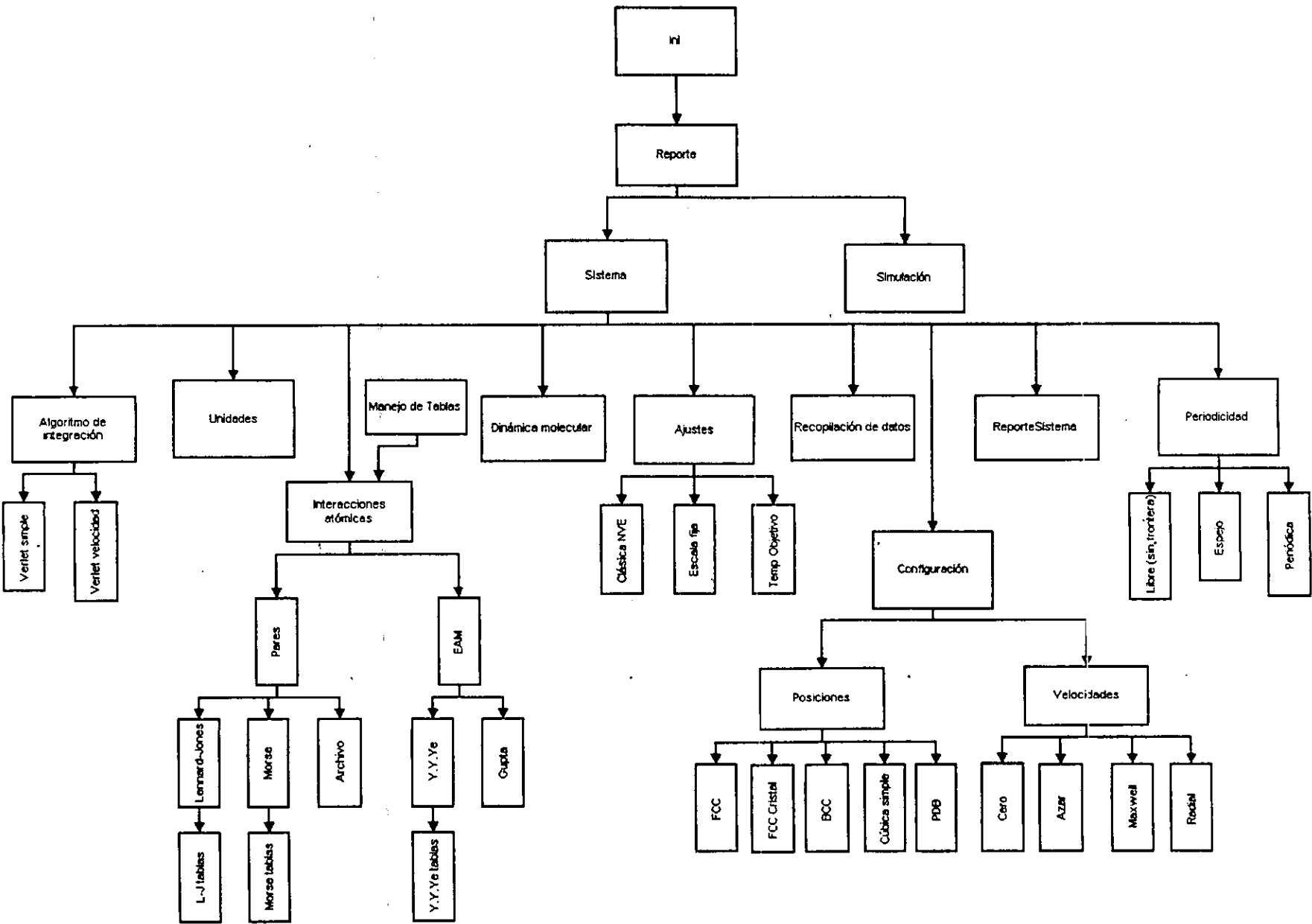


Figura 9-1 Diagrama de relaciones de derivación de clases (jerarquías).
DISEÑO Y FUNCIONES DEL PROGRAMA

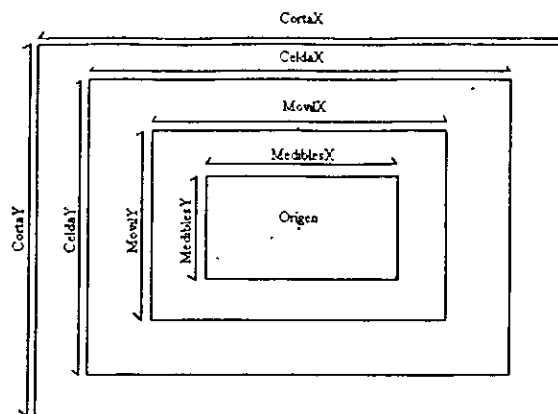


Figura 9-2 Diferentes celdas utilizadas para recortar/marcar/ordenar los átomos. Junto a cada una se muestra la instrucción correspondiente del archivo *.ini. Por claridad sólo muestro una celda bidimensional, pero el proceso ocurre en tres dimensiones.

4. cCero - Asigna todas las velocidades iniciales iguales a cero.
5. cClasicaNVE - Función derivada de cAjustes en donde se describe en forma explícita una función de ajuste vacía que permite al sistema evolucionar en forma libre o "natural" con energía constante.
6. cConfiguracion - Clase base que contiene las funciones virtuales encargadas de llenar el cluster y elegir las propiedades de los átomos, como si son medibles y/o móviles, recortar el cluster y reordenarlo.
7. cCubicaSimple - Llena el cluster con átomos en una configuración cúbica simple celda por celda.
8. cDinamicaMolecular - Contiene la función de evolución. Esta función organiza las llamadas a las funciones de medición, avance y reporte de otras clases.
9. cEAM - Definición explícita de la función que calcula la energía, fuerzas y virial para potenciales tipo EAM. Requiere de las funciones específicas a cada tipo de potencial, por ejemplo la función de densidad electrónica, la energía de embebimiento y el potencial de pares.
10. cEspejo - Definición explícita de las funciones de periodicidad en posiciones y distancias para sistemas con fronteras de espejo.
11. cFCC - Llena el cluster con átomos en una configuración cristalina FCC. Los átomos se colocan a partir de un átomo central, radialmente en capas, llenando cada orden de vecinos antes de pasar al siguiente.
12. cFCCcristal - Llena el cluster con un átomo en una configuración FCC cúbica. El número de átomos debe ser divisible por 4 porque es el número de átomos en una celda unidad. El cristal se construye celda por celda.
13. cGupta - Definición explícita de las funciones para el potencial Gupta de tipo EAM. Aún no implementada.
14. cIni - Es la clase más general dentro del árbol de jerarquías. Contiene las variables estáticas generales legibles del archivo ini, así como las funciones necesarias para leer este archivo. También incluye funciones para dar salida al ini tanto a pantalla como a un archivo. Aquí se definen también los valores máximos de número de átomos y vecinos así como de los elementos en las tablas y en las distribuciones. El algoritmo encargado de leer el archivo ini verifica la existencia de los valores buscados, así como problemas relacionados a la lectura.
15. cIntegracion - Definición virtual de las funciones que preparan la integración de un paso y que lo realizan usando las funciones de avance de un paso en posiciones y en velocidades declaradas aquí virtualmente.
16. cInteracciones - Contiene las funciones virtuales de cálculo de energía y fuerza en sus formas más generales. Contiene las funciones para actualizar las listas de vecinos y mostrarlas al archivo de bitácora. Aquí se encuentra la función que muestra un cuadro de la película y la que genera el archivo con las tablas del potencial y fuerza (de pares).

17. cLennardJones - Definición explícita de Er y Fr en forma analítica.
18. cLibre - Definición explícita de las funciones de periodicidad en posiciones y distancias para sistemas libres.
19. cLJ_tabla - Definición explícita de Er y Fr en forma de tablas que se construyen usando las formas analíticas de su clase base cLennardJones.
20. cManejoTablas - En esta clase defino la estructura necesaria para manejar tablas, con las variables necesarias para su aprovechamiento. He encontrado más fácil implementar los algoritmos para colocar y extraer datos de tablas en cada una de las clases que los necesitan. .
21. cMaxwell - Configura las velocidades de tal manera que cada componente pertenece a una distribución de Maxwell.
22. cMorse - Definición explícita de Er y Fr en forma analítica.
23. cMorse_tabla - Definición explícita de Er y Fr en forma de tablas que se construyen usando las formas analíticas de su clase base cMorse.
24. cPares - La función explícita que calcula la energía, fuerzas y el virial para potenciales de pares. Aquí se definen virtualmente las funciones Er y Fr que corresponden a la energía y fuerza de pares respectivamente.
25. cPares_Archivo - Definición explícita de Er y Fr, pero usando tablas llenadas con valores leídos de archivo. Aquí están las funciones y variables necesarias para leer los datos del archivo, interpolarlos y suavizarlos.
26. cPDB - Las funciones para leer un archivo PDB del disco.
27. cPeriodica - Definición explícita de las funciones de periodicidad en posiciones y distancias para fronteras periódicas.
28. cPeriodicidad - Define en forma virtual las funciones de periodicidad en posiciones, distancias e inicial. Aquí están las funciones que contienen los vectores que definen las celdas imagen y que permiten aprovechar la periodicidad.
29. cPosiciones - Clase base de los diferentes tipos de objetos que generan las posiciones (funciones FillCluster). Incluye las funciones para grabar una configuración en un archivo de formato PDB (común a todas las clases que generan configuraciones de posiciones).
30. cRadiales - Configura las velocidades tales que son idénticas en magnitud y dirección a la posición (después son escaladas para asignar la temperatura inicial).
31. cRecolectarDatos - Se encarga de recolectar datos del sistema, por ejemplo distribuciones, estadísticas y energía cinética. Conociendo los datos realiza muchas de las mediciones y puede verificar si es necesario realizar funciones de cReporte Sistema, como grabar configuraciones o estadísticas.
32. cReporte - Variables y funciones necesarias para dar salida. Para mantenerla general es necesario que sea una clase base de cSistema, sin embargo para poder utilizar variables de cSistema debe ser derivada. Por ello decidí separar las funciones según sus necesidades en cReporte y cReporteSistema.
33. cReporteSistema - Variables y funciones para dar salida a lo derivado de sistema. Utiliza funciones y variables de cReporte. Desde aquí se graban las configuraciones y demás propiedades del sistema.
34. cSimulacion - Funciones para instanciar los objetos adecuados, seleccionandolos de los posibles. La clase contiene funciones para preparar, equilibrar, simular y terminar, ya sea generales o de una etapa específica. Es el control maestro que pasa los apuntadores a cada objeto a los demás que los necesitan. Contiene la función para hacer de una cierta etapa n la etapa simulada.
35. cSistema - Incluye variables del sistema: distribuciones, variables de estado (energía, temperatura, densidad, etc...) y los átomos. También incluye las estructuras para estos tipos de variables.
36. cTobjetivo - Función derivada de cAjustes en donde se describe en forma explícita una función de ajuste que utiliza las funciones de su clase base para simular un sistema en el que las velocidades se escalan por el factor necesario para forzar en el sistema una cierta temperatura objetivo.

37. cUnidades - Funciones que permiten convertir unidades internas o reducidas a unidades reales y vice versa. No utilizadas pero listas.
38. cVelocidades - Clase base de los diferentes tipos de objetos que generan las velocidades.
39. cVerlet - Definición explícita de las funciones virtuales de avance de un paso definidas en cIntegración.
40. cVerletVelocidad - Definición explícita de las funciones virtuales de avance de un paso definidas en cIntegración.
41. cVEscaladaFijo - Función derivada de cAjustes en donde se describe en forma explícita una función de ajuste que utiliza las funciones de su clase base para simular un sistema en el que las velocidades se escalan por un factor fijo.
42. cYYYe - Definición explícita de las funciones para el potencial YYYe de tipo EAM.
43. cYYYe_tablas - Definición explícita de las funciones necesarias para un potencial YYYe de tipo EAM usadas por medio de tablas llenadas con las funciones analíticas de su clase base cYYYe.
44. Main - Controla las funciones de la clase cSimulación para ordenar la creación, preparación y uso de objetos. Allí se hace notar las diferencias entre las etapas. Decide que etapas se llevan a cabo y cuales no.
45. ManejoMemoria - Funciones que revisan si se ha liberado toda la memoria asignada por el programa en forma dinámica.

9.5 Los átomos

En una simulación de dinámica molecular se resuelven las ecuaciones que describen la evolución temporal de un sistema atómico. Por lo tanto es necesario mantener en algún lado la información de los átomos. En C++ una buena manera es declarar una estructura (por ejemplo de tipo Atomo) y luego reservar en memoria un arreglo de estructuras Atomo. Así lo hicimos. Para referirme a este conjunto de átomos utilizo la palabra *cluster*.

El arreglo de estructuras tipo Atomo se encuentra en la clase cSistema, donde también se encuentra la definición de la estructura Atomo.

La información almacenada de cada uno de los átomos es la siguiente:

- Número - El número del átomo, útil para cuando se lee una configuración de un archivo PDB.
- x, y, z - Las coordenadas cartesianas de su posición.
- r - Magnitud de su posición.
- vx, vy, vz - Proyecciones de la velocidad sobre los ejes cartesianos.
- fx, fy, fz - Proyecciones de la fuerza sobre los ejes cartesianos.
- OldFx, OldFy, OldFz - Proyecciones de la fuerza del instante anterior sobre los ejes cartesianos.
- Symbol - Símbolo químico del elemento del átomo (vg. Au)
- Vecinos[MAX_VECINOS] - Un arreglo de tamaño MAX_VECINOS de estructuras vecino. Cada estructura vecino incluye un apuntador al átomo del que se es vecino y un entero que representa el número de imagen en que es vecino.
- nVecinos - Número de vecinos.
- Movil - Si el átomo es móvil o no.
- Medible - Si el átomo debe ser usado para algunas mediciones o no.
- Grupo - El grupo al que pertenece el átomo, hasta ahora sólo usado para seleccionar los átomos mostrados en la película.

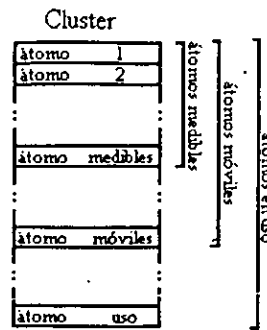


Figura 9-3 Orden de los átomos dentro del arreglo *cluster*.

- Ro - La densidad electrónica en este átomo. Para potenciales de tipo EAM.
- Bx, By, Bz. - Variables que almacenan proyecciones de la densidad. Para potenciales de tipo EAM.
- X0, Y0, Z0 - Posición del átomo en la última actualización de vecinos. Para decidir cuando debe actualizarse nuevamente.
- Además estos valores los elementos de un arreglo son identificables por su posición en el arreglo. Sin embargo el hecho de que en C los arreglos comienzan con el índice 0 puede complicarle la vida a un lector no atento, pues por lo general se acostumbra numerar las cosas a partir de 1.

9.5.1 Selección y marcado

De lo expuesto en las secciones anteriores queda claro que no siempre queremos hacer lo mismo con todos los átomos, por ejemplo, algunos son móviles y otros no. En lugar de repetir las decisiones acerca de la movilidad o inmovilidad de un átomo decidimos marcarlo (almacenar en la estructura Atomo) una bandera que indique el tipo de movilidad del átomo. Estas banderas se inicializan y son cambiadas conforme se necesite. Dado que es muy raro que una bandera deba cambiar de valor durante la simulación se gana mucho en eficiencia.

Una vez generadas las posiciones atómicas (o leídas de un archivo) es momento de revisar el cluster y adaptarlo a las necesidades específicas de la simulación. El programa procede a hacer estas modificaciones del cluster justo después de su creación. Todas las modificaciones son opcionales, es decir, podemos elegir llevarla a cabo o no. Nuestras decisiones las informamos al programa por medio del archivo de instrucciones (archivo ini).

En el diagrama de preparación general de la figura 9-5 se muestra el orden de las modificaciones hechas al cluster. A continuación describo los efectos de cada una de ellas:

1. Centrar Cluster - Se traslada el cluster completo de tal manera que el centro de masa coincida con el origen de coordenadas.
2. Recortar Cluster - Se eliminan todos los átomos que no estén contenidos en el paralelepipedo rectángulo centrado en el origen con lados CortaX, CortaY y CortaZ (indicados en el archivo de instrucciones). Un átomo *i* es borrado del cluster copiando todos los átomos desde *i*+1 hasta el último sobre el espacio asignado al anterior. Finalmente se disminuye el número de átomos usados (ATOMIOS_USO).
3. Periodicidad inicial - Se aplican la periodicidad definida en el archivo de instrucciones y se utilizan las imágenes requeridas para revisar que no se vayan a tener problemas más adelante. Para un sistema libre no se hace nada, pero para sistemas con periodicidad de espejo o periódica es importante. Por ejemplo, para sistemas periódicos se comparan las imágenes de cada átomo con todos los demás, revisando que no haya dos átomos (o sus imágenes) a menor distancia que un valor definido en las instrucciones. En caso de encontrar algo así uno de los átomos se elimina (nuevamente recorriendo el resto de los átomos del cluster).
4. Elegir átomos móviles - El valor por defecto es que todos los átomos son móviles, pero también existe la posibilidad de mantener algunos fijos en su posición inicial. La función que elige los átomos móviles es similar

a la que recorta el cluster, declarando móviles los átomos contenidos en un paralelepípedo rectángulo centrado en el origen de lados MovilX, MovilY y MovilZ (del archivo de instrucciones), pero en lugar de eliminar un átomo al declararlo inmóvil, los átomos son acomodados de tal manera que los átomos móviles son colocados al comienzo del cluster, hasta llegar a ATOMOS_MOVILES. De allí en adelante y hasta llegar al total de ATOMOS_USO todos los demás son inmóviles.

5. Asignar velocidades - Esta función asigna las velocidades a los átomos, pero solamente a los que han sido marcados como móviles.
6. Momento inicial cero - Se modifican las velocidades de todos los átomos, de tal manera que no haya una deriva del sistema que no modifica en nada su comportamiento global y hace más complicado el análisis.
7. Elegir átomos medibles - En forma análoga a como se han marcado y ordenado los átomos que decidimos fueran móviles, se marcan y ordenan los átomos que consideramos medibles, es decir, de los que nos interesa obtener algunas mediciones especiales más adelante. La celda dentro de la cual se consideran los átomos como medibles tiene lados MediblesX, MediblesY y MediblesZ. Sólo vamos a necesitar considerar medibles átomos que son móviles, por lo que no se pierde el orden anterior conseguido en "Elegir átomos móviles", sino que se refina.
8. Fijar temperatura inicial - Se escalan las velocidades para que la temperatura inicial sea la indicada en el archivo de instrucciones.
9. Contar átomos visibles - Según el grupo al que pertenecen (ya sea por default o según son leídos de un archivo de configuración) y según un parámetro indicado en las instrucciones, los átomos serán grabados o no en la película. Por razones operativas determinadas por el formato de salida de películas, es necesario saber desde el comienzo el número de átomos que serán visibles. Aún no se aprovecha a fondo esta posibilidad, pero la intención es que si más adelante encontramos una razón para sólo grabar algunos de los átomos, podamos definir una función que los marque.

En la figura 9-2 muestro la forma en que se utilizan las diferentes celdas para recortar/marcar/ordenar átomos. Por claridad muestro las celdas en esas proporciones, pero no necesariamente es el caso. Sin embargo, si existen ciertas restricciones al tamaño de cada una de las celdas, por ejemplo. La celda de átomos medibles debe estar contenida (o a lo más ser igual) en la celda de átomos medibles, de lo contrario habría átomos medibles-inmóviles. un caso no permitido por el programa. Para más información sobre los requisitos de cada una de las celdas sugiero las secciones donde se justifica el uso de cada una, así como la sección 9.8 que describe los elementos del archivo de instrucciones.

Dado que diferentes métodos para marcar propiedades en los átomos los reacomodan dentro del arreglo "cluster". En la figura 9-3 muestro el arreglo final resultado de aplicar los métodos de selección de móviles y de selección de medibles. En la figura se observa como los átomos medibles son un subconjunto de los móviles. Por claridad para el lector el arreglo comienza con el índice 1, por lo que el átomo "medibles" es también medible. En C los arreglos comienzan con índice 0, por lo que el átomo "medibles" ya no es medible.

9.6 Flujo

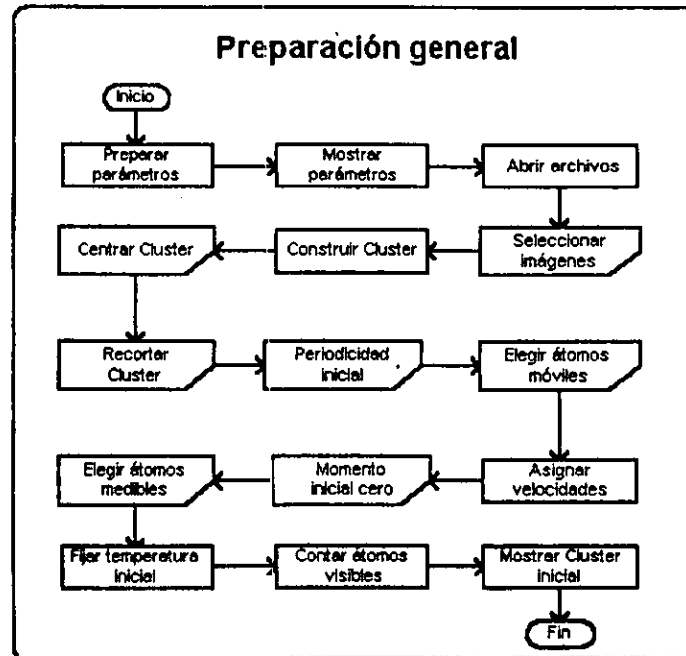
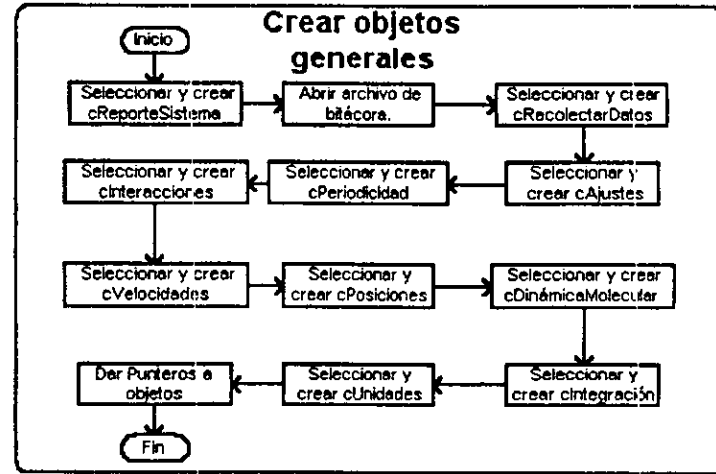
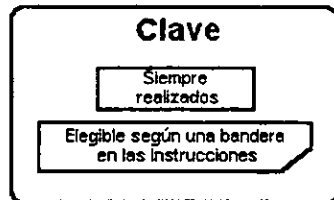
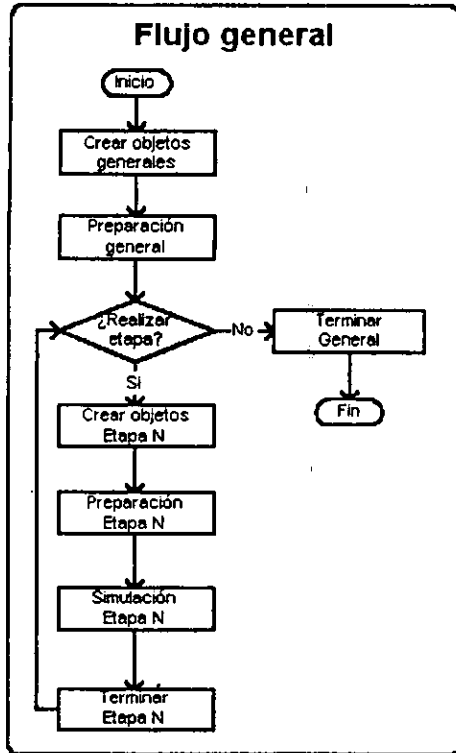
9.6.1 Etapas

Como describo en la sección 6.4 es común que sea necesario llevar al sistema a un estado de equilibrio antes de poder realizar una simulación de la cual obtener mediciones útiles. Esto significa al menos dos evoluciones consecutivas diferentes, cada una con sus propios objetivos y parámetros.

Hay diferentes utilizar un programa para hacer evoluciones consecutivas. Una es generar archivos al final de cada evolución con la información completa, de tal manera que puedan ser utilizados como fuente. A pesar de tener máxima versatilidad este método tiene inconvenientes porque implica la capacidad de grabar, almacenar y leer una enorme cantidad de información, la mayor parte inútil. La información sería tanta que sólo el mismo programa que la genero sería capaz de comprenderla, por lo que nuevamente comenzamos a perder versatilidad.

La solución es permitir al programa hacer evoluciones consecutivas, aprovechando la información almacenada en memoria. Nosotros elegimos esta alternativa. Nuestro programa lo diseñamos de manera que recibiera instrucciones

Figura 9-4 Diagramas de flujo parte 1.



para un máximo de cinco etapas diferentes. Cada una de las etapas tiene su propio tipo de dinámica, sus propias cantidades de ciclos ciegos y ciclos generales, su propio intervalo de tiempo δt , y los parámetros necesarios para el tipo de dinámica elegida. Al final de cada etapa se crea un archivo pdb con la configuración, para poder observar los efectos de la simulación y que puede servir como configuración de inicio para futuras simulaciones. Además es posible decidir realizar o no realizar cada una de las etapas dependiendo del valor de una bandera en las instrucciones.

Las funciones que usa cada una de las etapas son exactamente las mismas, las únicas diferencias provienen del tipo de dinámica elegida en cada etapa. Para que sea posible hacer esto las funciones tienen que utilizar variables generales. Para esto el programa tiene variables generales (llamadas "de uso") adicionales a las de cada etapa (llamadas EtapaN). Una vez que la bandera adecuada ha indicado que una etapa se ha de realizar los parámetros de la etapa son copiados a los parámetros en uso. De esta forma se continúa y en cuanto a los parámetros el haber pasado de una etapa a otra es transparente para el programa.

Respecto a los objetos para cada una de las etapas es necesario construir (instanciar) los objetos específicos de cada etapa, destruyendo los que se han vuelto innecesarios. Cuando se han cambiado los objetos hay que actualizar la información que cada objeto tiene respecto a los demás. Agregamos también funciones que permiten a cada etapa realizar los preparativos y los trabajos de terminación específicos para su etapa. De esta forma el programa puede ejecutarse y realizar 5 etapas una tras otra o no ejecutar ninguna.

9.6.2 Diagramas de flujo

"Cuidado con el programa anterior, solo he demostrado que no tiene errores, no lo he probado".
Donald Knuth [Han99].

Los diagramas de flujo las figuras 9-4, 9-5 y 9-6 representan los pasos principales del programa. Desde el nivel más general de "Flujo general", hasta el más específico de "Cálculo de la energía potencial y las fuerzas". Cada uno de los diagramas tiene un óvalo para indicar el lugar de inicio y otro para el fin. Los rectángulos representan funciones, si la esquina inferior derecha está recortada son funciones que pueden ser efectuadas o no dependiendo de una bandera en el archivo de instrucciones. Los rombos, como es costumbre, representan decisiones. A continuación describo brevemente acerca de cada uno de los diagramas.

- **Flujo general** - (figura 9-4) Nivel más general. Después de leer las instrucciones del archivo ini se crean los objetos que existirán a lo largo de toda la ejecución. Creados los objetos se prepara al sistema. Se seleccionan las etapas a realizarse, cada una consistiendo de la creación de los objetos particulares, la preparación, la simulación y la terminación. Finalmente la terminación general.
- **Crear objetos generales** - (figura 9-4) Primero se crea un archivo para la bitácora, para poder grabar en ella todos los procesos siguientes. Se crean los objetos que existen durante toda la ejecución, seleccionándose uno dentro de cada una de las ramas principales de la jerarquía de clases. Una vez creados los objetos se reparten sus direcciones para que puedan interactuar.
- **Preparación general** - (figura 9-4) La preparación de parámetros calcula algunos valores que sin ser obtenidos del archivo de instrucciones son directamente derivables. También revisa la congruencia de algunos valores. por ejemplo del número de átomos con el tipo de configuración de posiciones seleccionada. Preparados los parámetros se muestran (a pantalla y al archivo de bitácora), se crean los archivos para todo lo que no sea bitácora (ya existente). De ser necesario se seleccionan las celdas imagen resultantes de las simetrías indicadas. Se construye el cluster (se llenan los espacios de posiciones). De ser necesario se centra el cluster llevando el centro de masa al origen. De ser necesario se recorta el cluster, eliminando los átomos fuera de la celda de corte. De ser necesario se verifica la periodicidad inicial, dependiendo del tamaño de la celda de periodicidad y de las simetrías seleccionadas. En caso de indicarlo el mi se eligen los átomos que serán móviles y a todos los que lo sean se les asignan velocidades. Según las instrucciones se pueden modificar las velocidades para hacer el momento lineal inicial cero. De ser necesario se eligen los átomos medibles (sobre los que se harán mediciones). Se fija la temperatura inicial, se cuentan los átomos visibles (los que aparecerán en la película) y se genera un archivo con el cluster inicial.
- **Crear objetos etapa N** - (figura 9-5) Decidida la realización de una etapa N se llega a este diagrama. Comienza por copiar los parámetros de la etapa en curso (N) a los espacios de los parámetros en uso. Se destruyen los objetos innecesarios anteriores y se seleccionan y construyen los objetos específicos a esta etapa. Finalmente se reparten nuevamente las direcciones de los objetos existentes para actualizar las ya obsoletas.

Figura 9-5 Diagramas de flujo parte 2.

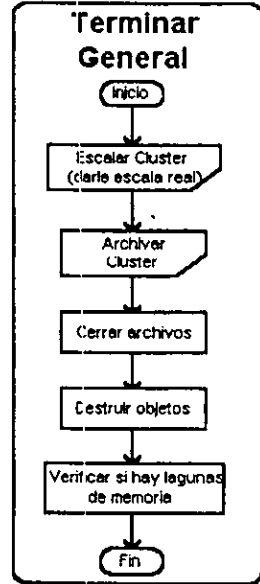
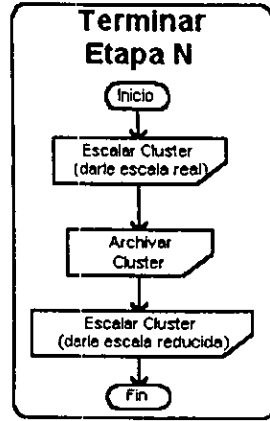
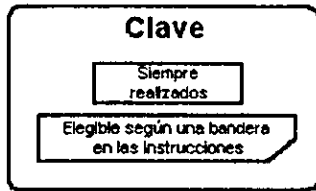
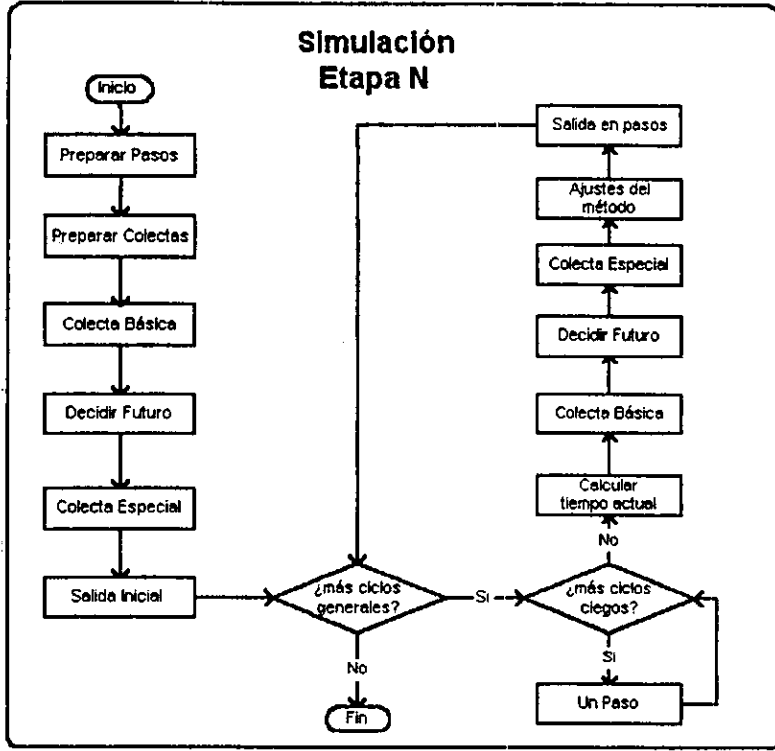
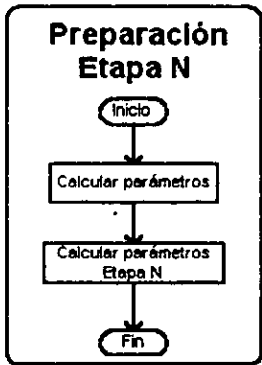
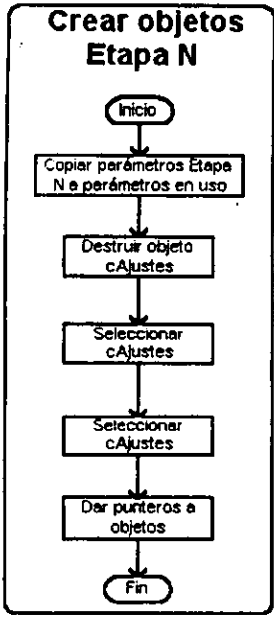
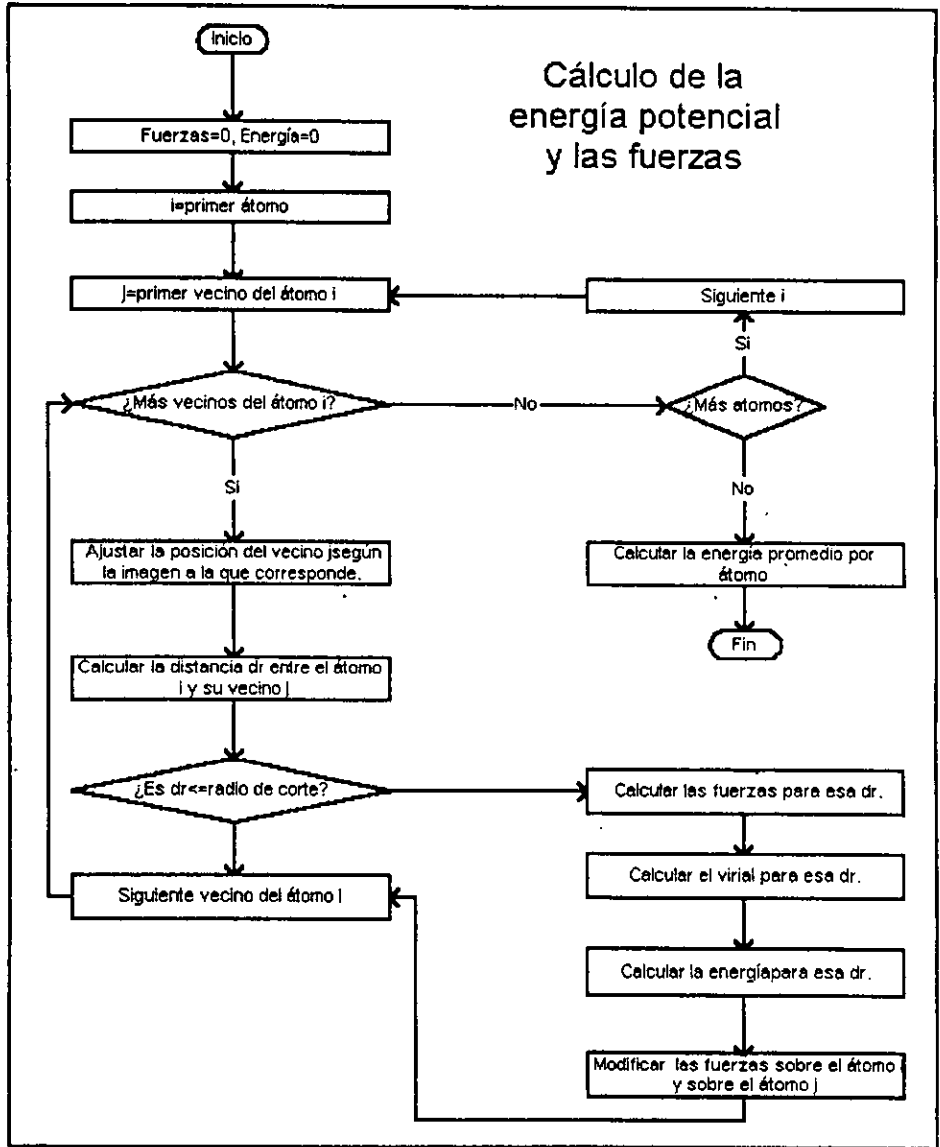
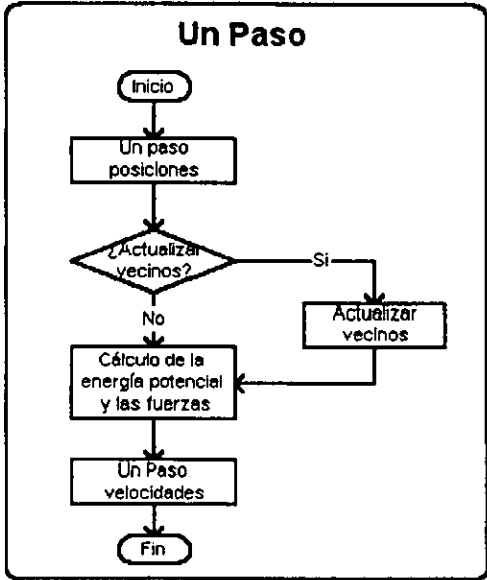


Figura 9-6 Diagramas de flujo parte 3.



- **Preparación etapa N** - (figura 9-5) La preparación de una etapa N depende del tipo de evolución que signifique, por ello es necesario separarla en sus partes general y específica. Se calculan los parámetros generales y luego los parámetros específicos.
- **Simulación etapa N** - (figura 9-5) Este diagrama representa el proceso de evolución. Se supone un sistema ya listo, pero es necesario hacer algunas mediciones antes de comenzar, para ellos se preparan los pasos (proceso que es diferente para cada algoritmo de integración), se preparan las variables que pueden ser medidas y recolectadas para calcular estadísticas. Se hace la colecta básica (los valores del sistema que siempre queremos conocer). En "decidir futuro" se seleccionan el resto de las acciones a realizar antes de comenzar la evolución, incluyendo cálculos, distribuciones, salidas a archivo y otros. Durante "Colecta Especial" se llevan a cabo las acciones elegidas. En "Salida inicial" se muestran los resultados, para poder conocer el estado previo al inicio de la evolución.

Se inician entonces los ciclos generales. Cada uno consiste de ciclos ciegos y de las mediciones y alteraciones del sistema. En cada ciclo ciego se avanza al sistema un paso en el tiempo. Transcurridos ciclos ciegos se efectúa primero la colecta básica, nuevamente se deciden el resto de las acciones a seguir, dependiendo del tipo de simulación que se realice y del tiempo actual. En la colecta especial se realizan las acciones. Se ajusta el sistema para que su evolución sea la indicada por las instrucciones y se da salida a los resultados. En cuanto se han llevado a cabo ciclos generales se termina la evolución. Notese que los ajustes se llevan a cabo después de efectuar los cálculos, pues estos son necesarios para decidir los ajustes a realizar, y la salida se hace después de los ajustes, para mostrar un sistema tal y como comienza el siguiente ciclo general (o termina la evolución).

- **Terminar etapa N** - (figura 9-5) Para poder dar por terminada una etapa N pueden ser necesarios diversos ajustes. Para los tipos de etapas que hemos implementado lo único necesario es generar un archivo pdb con la configuración del cluster. Para que esta configuración sea comparable con la inicial y con los pdb's estándar se escala el cluster antes de mostrarlo. Finalmente se vuelve a escalar para regresarlo a las unidades internas y poder continuar.
- **Terminar general** - (figura 9-5) La terminación general no sólo graba la configuración como la terminación de una etapa N, sino que además cierra los archivos que se usaron durante la simulación, destruye los objetos (inclusive algunos que han dejado de ser idénticos a los iniciales, pero que fueron creados por etapas). Justo antes de terminar se verifica que la memoria requerida por el programa en forma dinámica haya sido liberada (por ejemplo que todos los objetos creados hayan sido destruidos).
- **Un paso** - (figura 9-6) Es el elemento básico de la dinámica molecular. En un paso el sistema avanza dt en el tiempo. Dependiendo del algoritmo de integración un paso puede ser ligeramente diferente. En la figura 9-6 se muestra como ejemplo un paso del algoritmo de Verlet Velocidad. Consiste en avanzar las posiciones de las partículas según las velocidades del instante anterior. Con las nuevas posiciones se decide si es necesario actualizar las listas de vecinos de cada átomo. De ser necesario se hace y se continua. Con las nuevas posiciones y listas de vecinos válidas se actualizan las fuerzas, la energía, y el virial. Finalmente se utilizan las nuevas fuerzas para actualizar las velocidades.
- **Cálculo de la energía potencial y las fuerzas** - (figura 9-6) Es el corazón de un programa de dinámica molecular. Una ineficiencia en este algoritmo puede tener consecuencias catastróficas en el tiempo de simulación. Comenzamos por inicializar las fuerzas y la energía para todos los átomos. Se elige un átomo y se recorren todos sus vecinos. Para cada uno de ellos se hacen primero las correcciones a su posición para representar los efectos de la simetría (la celda imagen en que se encuentra), con estas correcciones se calcula la distancia (en x , y , z y en línea recta) entre el átomo i y la imagen k del vecino j . Si la distancia en línea recta es mayor al radio de corte se pasa al siguiente vecino. Si es menor o igual se calculan las fuerzas entre ambos, el virial y la energía potencial asociada. Al calcular las fuerzas sobre el átomo i la 3a ley de Newton nos dice que las fuerzas sobre el vecino j serán las opuestas, así que se las asignamos. Con ello podemos pasar al siguiente vecino. Terminados los vecinos del átomo i pasamos al siguiente átomo y al concluir con todos los átomos sólo resta dividir la energía potencial entre el número de átomos para conservar la energía promedio.

9.7 Manejo de unidades

9.7.1 Internas

En la sección 6.1 describo las ventajas del manejo de unidades reducidas. En esta sección describo la forma en que se aplico esta teoría en el programa generado con esta tesis.

Las ecuaciones de cada potencial que el programa conoce en forma analítica estan simplificadas en la forma más adecuada a cada uno, por lo que las unidades reducidas son ligeramente diferentes en cada caso. Preferí hacer esto y complicar un poco el uso de las unidades a cambio de simplificar la escritura/entendimiento del programa, facilitando así su futura modificación. Un gran aparato de modificación de unidades hubiera oscurecido algunos aspectos fundamentales del funcionamiento del programa y dificultado su adaptación a nuevos potenciales. El efecto de esta "variabilidad" de las unidades reducidas con el potencial es muy pequeño, por lo que tampoco significa problemas importantes.

Tanto para los potenciales de pares analíticos como para los leídos de un archivo se hicieron las siguientes suposiciones:

1. Las unidades de distancia son unidades reducidas. Es decir, la energía mínima ocurre para $r_{E=\min} = 1$.
2. Las unidades de masa son reducidas, con $m = 1$ (la masa de un átomo).

Sin embargo hay diferencias entre la suposición en la energía para ambas fuentes de potenciales.

Para potenciales analíticos: Las unidades de energía son reducidas, $E_{\min} = -1$ (la energía mínima).

Para potenciales de archivo: Las unidades de energía estan en eV, y son del orden de la unidad.

Respecto al tamaño semejante a la unidad de los potenciales de archivo encontramos que para todos los potenciales de pares de interés en esta tesis pudimos construir archivos de este tipo. Es trabajo del usuario prepararlos.

Esta diferencia entre potenciales analíticos y de archivo (tabla) significa que no podemos preparar y analizar de la misma manera sistemas en ambas circunstancias. Por supuesto esto significa un trabajo extra para el usuario, pero ha sido hecho de esta forma para maximizar la adaptabilidad del programa, y reducir a su vez el tiempo de preparación de los archivos de potencial.

Suponer que tanto en las funciones analíticas como en los archivos la distancia de equilibrio de los primeros vecinos es a $r_{E=\min} = 1$ implica escalar las distancias del cluster leído/generado antes de poder utilizarlas. El programa es capaz de hacer esto, pero puede ser activado/desactivado con una banera en el archivo de instrucciones (*.ini). Si se ha modificado la escala del cluster el programa realiza el escalamiento inverso antes de grabar los archivos de configuración en el disco (PDB's), de tal manera que puedan ser comparados directamente con los archivos de entrada (sección 9.9).

9.7.2 Conversion de unidades internas \longleftrightarrow externas

Para poder hacer las conversiones de unidades reales a reducidas y vice versa el programa recibe tres parámetros del archivo de instrucciones:

1. La distancia de interacción de mínima energía $r_{E=\min}$. Para muchos casos la asumo idéntica a la distancia de equilibrio de primeros vecinos. Esta debe estar en Angstroms.
2. La energía mínima E_{\min} . Esta debe estar en Joules/mol. No necesariamente es la unidad en que conocemos las características del material, sin embargo es necesario escoger una unidad estándar y por lo general la conversión a esta unidad es directa y simple.
3. La masa de un átomo. En Kilogramos/mol.

Una vez conocidos estos parámetros para un material quedan determinadas las unidades de longitud, energía y masa. Con combinaciones de ellas es posible determinar el resto de las unidades. A continuación describo las reglas de conversión de unidades reducidas a unidades reales necesarias cuando se utiliza un potencial de Lennard-Jones analítico reducido. Cuando lo que se utiliza un potencial de archivo la unica diferencia es simplificar la regla de

Parámetros del archivo ini	Ecuación de conversión	Ejemplo L-J y Argón
Masa	$m^* = 1$	$M = 0.0402897 \text{ kg}\cdot\text{mol}^{-1}$
Longitud	$r^* = r/\sigma$	$r_{\mathcal{E}=\min} = 3.816371 \text{ \AA}$
Energía	$\mathcal{E}^* = \mathcal{E}/\epsilon$	$\mathcal{E}_{\min} = 997.68 \text{ J}\cdot\text{mol}^{-1}$
Unidad		
Longitud	$r = r^* \sigma = r^* \frac{r_{\mathcal{E}=\min}}{2^{1/6}}$	$r = r^* * 3.4 \times 10^{-10} \text{ m}$
Energía	$\mathcal{E} = \mathcal{E}^* \mathcal{E}_{\min}$	$\mathcal{E} = \mathcal{E}^* \times 997.68 \text{ J}\cdot\text{mol}^{-1}$
Masa	$m = m^* M$	$m = m^* \times 0.0402897 \text{ Kg}\cdot\text{mol}$
Tiempo	$t = t^* \sigma \sqrt{\frac{M}{\mathcal{E}_{\min}}}$	$t = t^* \times 2.16063 \times 10^{-12} \text{ s}$
Densidad	$\rho = \rho^* \sigma^3 = \rho^* \frac{\sqrt{2}}{2} r_{\mathcal{E}=\min}$	$\rho = \rho^* \times 1.81238 \text{ g}\cdot\text{cm}^{-3}$
Temperatura	$T = T^* \frac{\mathcal{E}_{\min}}{8.314}$	$T = T^* \times 120 \text{ }^\circ\text{K}$
Presión	$P = P^* \frac{\mathcal{E}_{\min}}{\sigma}$	$P = P^* \times 2.3724 \text{ N}\cdot\text{m}^{-2}\cdot\text{átomo}$

Tabla 9.1 Conversión de unidades.

transformación de la energía. La siguiente tabla de conversiones debe ser comparada con la tabla 6.1 de la sección 6.1 para observar la aplicación de las fórmulas allá mostradas.

Hay que hacer notar que además de conocer las escalas de distancia, energía y masa se conocen también diferentes factores de conversión y constantes físicas. Ellas las incluyo en el apéndice C.

Dado que no todas las unidades son siempre modificadas a una versión reducida (por ejemplo la energía en los potenciales de pares de archivo), he preferido llamarles *unidades internas*, en lugar de reducidas. En este caso hay que tener en mente una diferencia en los resultados presentados por el programa y en lo que espera de instrucciones: La temperatura no está en las mismas unidades para todos los potenciales. Si los potenciales están en unidades reducidas (ie. dimos el factor de conversión de la energía y $\mathcal{E}_{\min}^* = -1$) el programa recibe y devuelve energías y temperaturas reducidas. Si el potencial es de un archivo y no usamos energías reducidas el programa recibe y devuelve energías cinéticas y temperaturas en la unidad en que se entrego la energía (usualmente eV como dije antes).

El programa cuenta con las funciones que hacen la transformación de unidades reducidas a reales y vice versa, una vez que se han dado los parámetros adecuados en el archivo de instrucciones. Sin embargo no aprovecho estas funciones para que el resto de las instrucciones y las salidas estén en unidades reales porque esto eliminaría muchas de las ventajas de utilizar unidades reducidas. Se aumentan los trabajos de preparación y de análisis, pero se recupera en simplicidad y versatilidad. Para evitar las diferencias al usar potenciales de pares analíticos o de archivo se podría escalar el archivo recibido hasta que $\mathcal{E}_{\min}^* = 1$, pero esto significaría una de dos opciones: a) Sustituir un trabajo del usuario por otro. b) Un algoritmo que encuentre el mínimo y haga el escalamiento, lo que ha mostrado ser poco preciso. Sin embargo a futuro pienso implementar la segunda opción.

9.8 Instrucciones, el archivo ini

“Probe el programa de todas las maneras que se me ocurrieron. Lo force para exponer sus debilidades. Lo ejecute para estrellas de masa grande y pequeña, para estrellas nacidas extremadamente calientes y aquellas nacidas relativamente frías. Lo ejecute asumiendo que las corrientes de superfluido bajo la corteza no existían. No porque quisiera saber la respuesta, sino porque había desarrollado una sensación intuitiva de la respuesta en este caso particular. Finalmente conseguí una corrida en la que la computadora mostro que la temperatura de un pulsar es menor que el cero absoluto. Había encontrado un error. Perseguí el error y lo corregí. Ahora tengo una versión mejorada del programa a tal punto que no funciona en absoluto.”, George Greenstein en su libro “Frozen star: Of pulsars, Black Holes and the fate of stars [Han99].

Para programas con diferentes opciones es necesario especificar el tipo de simulación que se requiere en una ejecución específica, así como los archivos de entrada y salida específicos. La mejor opción es generar un archivo de terminación ini, que contenga todos los parámetros necesarios y que sea leído al comienzo de la ejecución. De esta forma nos evitamos la larga lista de parámetros repetidos que se tienen que dar a programas en los que es necesario dar entrada directamente desde el teclado a las opciones.

Usando un archivo ini podemos almacenar las opciones con que se ejecuto una simulación, para poder repetirla más tarde. El programa recibe de la línea de comando el nombre del archivo ini a usar (hay un valor por defecto), de esta forma podemos preparar listas completas de archivos ini, a ejecutarse sin necesidad de vigilancia.

Decidimos hacer un archivo ini complejo, es decir, que contenga tantos parámetros como sea posible modificar en el programa. Así no es necesario recompilar para modificar ninguna variable. La dificultad es que los archivos ini involucrados son grandes y complejos comparados con los de otros programas de simulación. La manera en que solucionamos este problema es escribiendo un programa de interfase que facilita la lectura, entendimiento y escritura de los archivos ini.

En el archivo de ini hay cuatro tipos diferentes de instrucciones:

1. Parámetros- Son valores necesarios en las ecuaciones dentro de los algoritmos. Por ejemplo el número de átomos, la temperatura inicial y el ancho de la celda de simulación. Por lo general son números reales, aunque pueden también ser enteros o cadenas de caracteres.
2. Banderas - Determinan si un proceso debe o no llevarse a cabo, son enteros y solo pueden tomar valores 0 ó 1. Por ejemplo si se debe o no hacer cero el momento lineal inicial o si se debe o no grabar una película.
3. Nombres - Son cadenas que contienen el nombre y dirección de archivos de entrada y/o salida. Por ejemplo el nombre del archivo en que se graban las distribuciones de velocidades.
4. Opciones - Determinan el tipo de proceso que se ha de llevar a cabo. Por ejemplo el tipo de interacción (potencial) o el tipo de configuración atómica.

La manera en que se indica el nombre y dirección del archivo de instrucciones (que no necesariamente tiene que tener terminación ini) es en la línea de comando al ejecutar el programa. Es el único parámetro que puede recibir el programa así que no hay que preocuparse por el orden. Si no se especifica ningún archivo de instrucciones el programa busca uno por defecto. Este puede ser cambiado solamente en el código fuente y por comodidad a la hora de programar y por la organización específica de mi computadora el valor por defecto es "E:\Molecular\inis\probando.ini".

Una vez elegido el tipo de simulación el programa usa las instrucciones adecuadas, lo que significa que no siempre son necesarias todas las instrucciones. En estos casos el programa simplemente ignora las instrucciones innecesarias. Esta permitido no incluirlas en el archivo *.ini, sin embargo su ausencia se hace notar en el archivo de bitácora.

9.8.1 Elementos del archivo ini

A continuación doy una lista de los elementos del archivo de intrucciones *.ini, tal y como se encuentran en un archivo típico. Pueden estar en cualquier otro orden. Las dos líneas diagonales indican comentarios que, por supuesto, no son interpretados. Los elementos van separados por una coma "," de su valor asignado. En itálicas presento un ejemplo de valor posible.

1. Nombres y direcciones de archivos de entrada.
 - (a) *PDB_In,E:\Molecular\PDBs\C60.pdb*. Nombre y dirección del archivo con las configuraciones a leer de disco. Con la información de este archivo se inicializa el cluster. No se obtienen de él más que las posiciones, el grupo y elemento, el resto de los parámetros se inicializa según las instrucciones. Sólo es necesario cuando Tipo_Config_posiciones=1.
 - (b) *Potencial_In,E:\Molecular\Potencial_In\AL_LJ.xy* - Nombre y dirección del archivo con el potencial de pares. Sólo es necesario cuando Tipo_Interacciones=2. Los formatos más comunes son *.xy y *.dat, la extensión no importa pero el programa espera que tenga dos columnas, una con r y otra con V(r).
2. Nombres y direcciones de archivos de salida.
 - (a) *Log_Out,E:\Molecular\logs\logout.txt* - Nombre y dirección del archivo de bitácora. Siempre es necesario un archivo de bitácora.
 - (b) *PDB_Out_Inicial,E:\Molecular\PDBs_Out\Inicio.pdb* - Nombre y dirección del archivo con la configuración inicial. Esta se graba a disco justo antes de comenzar la evolución, ya aplicados los mecanismos de selección. Siempre se crea este archivo.

- (c) PDB_Out_Final, *E:\Molecular\PDBs_Out\Salida.pdb* - Nombre y dirección del archivo con la configuración final. Esta se graba a disco justo antes de terminar la ejecución del programa, y antes de destruir el cluster. Sólo se crea cuando Flag_SaveEndCluster=1
- (d) Film_File, *E:\Molecular\films\pelicula.xyz* - Nombre y dirección del archivo con la película. El formato es adecuado para Re_View 1.0.
- (e) Raw_Out, *E:\Molecular\OutRaw\SalRaw.dat* - Nombre y dirección del archivo con las mediciones "crudas". El formato es de varias columnas. En orden desde la primera hasta la 5ª columna los datos de salida son: tiempo, energía total, energía cinética, magnitud del momento lineal total y presión. Sólo es necesario si Flag_Raw=1.
- (f) Cooked_Out, *E:\Molecular\OutCook\SalCook.dat* - Nombre y dirección del archivo con las mediciones "cocinadas". El formato es de varias columnas. En orden desde la primera hasta la 3ª los datos de salida son: tiempo, energía total promedio, varianza de la energía total. Sólo es necesario si Flag_Cooked=1.
- (g) DistPares_Out, *E:\Molecular\DistPar\DistPar.dat* - Nombre y dirección del archivo con la distribución de pares. El formato del archivo es r vs N(r) donde N(r) es el número de parejas de átomos a esa distancia. Las características como número resolución en la separación entre átomos y si esta distribución es instantánea o promediada en el tiempo están especificados por los parámetros de la sección de mediciones de distribuciones. Sólo es necesario si Flag_DistPares=1.
- (h) DistVelocidades_Out, *E:\Molecular\DistVel\DistVel.dat* - Equivalente al anterior, pero para la distribución de velocidades. Excepto a la hora de hacer las mediciones el proceso es equivalente a las distribuciones de pares. Sólo es necesario si Flag_DistVel=1.
- (i) Potencial_Out, *E:\Molecular\Ptcial_Out\potout.dat* - Nombre y dirección del archivo con el potencial de pares utilizado. La salida se construye con la misma información que se usa para hacer los cálculos internos, por ello este potencial permite estudiar y verificar el funcionamiento de los algoritmos para generar tablas, leer archivos, interpolar y demás funciones internas relacionadas con el manejo de potenciales. Sólo es necesario si Flag_Potencial=1.

3. Banderas de salidas.

- (a) Flag_Grabar_Imagenes_PDB, 1 - Bandera que determina si al grabar las configuraciones en los archivos PDB se graban también las celdas imagen definidas por la periodicidad. Tiene efecto sobre todos los archivos pdb generados. 0= No grabar imágenes. 1=Si grabar imágenes.
- (b) Flag_SaveEndCluster, 1 - Bandera que determina si se graba o no un pdb con la configuración final. En ocasiones no es necesario porque se ha grabado ya al final de la etapa correspondiente. 0=No grabar cluster. 1=Si grabar cluster.
- (c) Flag_Raw, 1 - Banera que determina si se graba o no el archivo con las mediciones "crudas". 0=No grabar archivo. 1=Si grabar archivo.
- (d) Flag_Cooked, 0 - Bandera que determina si se graba o no el archivo con las mediciones "cocidas". 0=No grabar archivo. 1=Si grabar archivo.
- (e) Flag_Cluster, 1 - Banera que determina si se muestra o no el cluster inicial en el archivo de bitácora. Este se muestra justo después de leerlo del pdb o generarlo, pero antes de cualquier modificación. Esto permite verificar que se ha leído correctamente. algo especialmente útil dado que el estándar de pdb no es respetado por la mayoría de los programas que lo usan (incluyendo este!). 0=No mostrar archivo. 1=Si mostrar archivo.

4. Referentes al archivo de película.

- (a) Flag_Film, 1 - Bandera que determina si se graba o no un archivo con la película. 0=No grabar archivo. 1=Si grabar archivo.
- (b) FilmPeriod, 300 - Período con el que se graban los cuadros de la película. Nótese que la verificación de si es momento de grabar un cuadro o no revisa el número de ciclos transcurridos hasta el momento módulo este período. por lo que es necesario que este número sea un múltiplo de ciclos_ciegos, de lo contrario no se grabara ningún cuadro.

- (c) Grupos_grabar,4 - Determina el máximo grupo del que se graban los átomos en la película. El grupo del átomo puede ser leído del pdb o inicializado por el programa. De esta forma podemos leer pdb's con periodicidad generados por este mismo programa, y/o mostrar únicamente algunos átomos en la película. Esto es importante porque ayuda a hacer más claras las animaciones y para no sobrepasar el pequeño límite de memoria permitido por Re_View. Uno de los usos comunes que nosotros le dimos es grabar solamente la película de los átomos de la frontera de grano.
- (d) Imagenes_Grabar_Film,1- Determina el número de imágenes que se graban en la película. No permite elegir cuales son las imágenes que se graban, pero sigue las simetrías según hayan sido definidas las combinaciones por las simetrías indicadas por Flag_Usar_Imagenes_X, Flag_Usar_Imagenes_Y y Flag_Usar_Imagenes_Z. Por ejemplo, Imagenes_Grabar_Film=1 graba solo la celda original, mientras que Imagenes_Grabar_Film=3 graba la celda original y dos celdas imagen sobre el eje x.

5. Referentes a la configuración de posiciones.

- (a) Tipo_Config_posiciones,3 - Opción para especificar el tipo de configuración de posiciones atómicas iniciales. El número de átomos es especificado por ATOMOS_USO para cuando el programa genera las posiciones y es el número de átomos en el archivo para configuraciones leídas de archivo pdb. 0= El programa genera un cluster con los átomos en las posiciones de un cristal FCC partiendo de un átomo en el origen y llenando las capas de vecinos hasta completar el número de átomos (llenando capas, como en una cebolla). 1=Se lee de un archivo PD. 2= Átomos en configuración cristalina cúbica simple. 3= Átomos en configuración FCC, pero a diferencia del tipo cebolla aquí se llenan cubos de átomos. 4=Átomos en configuración cristalina BCC. Algunas de estas configuraciones requieren que ATOMOS_USO sea un múltiplo del número de átomos en una celda unidad, en este caso el programa reduce ATOMOS_USO para que esto se cumpla.
- (b) ATOMOS_USO,130 - Número de átomos en el cluster. Sólo es necesario si Tipo_Config_posiciones≠1. porque en este caso la configuración se obtiene de un archivo.
- (c) Escala_Inicial,3.5 - Después de llenar las posiciones atómicas el cluster se escala por este parámetro para cambiar las unidades de distancia a unidades internas. Las posiciones se multiplican por este factor. Antes de la grabación de cualquier archivo pdb se aplica otro escalamiento por $\frac{1}{Escala_Inicial}$ para regresar a las unidades del cluster inicial y permitir comparaciones con una configuración leída de pdb. Una vez grabado el pdb se escala nuevamente para regresar a las unidades internas y continuar con la simulación. Sólo es necesario si Flag_Escala_Inicial=1.
- (d) Flag_Escala_Inicial,1 - Bandera que determina si se escala o no el cluster en los momentos indicados arriba. 0=No hacer escalamientos del cluster (equivalente pero menos eficiente que hacer Escala_Inicial=1). 1= Si hacer escalamientos del cluster.
- (e) CELDAX,15.8 - Ancho en x de la celda original centrada en el origen en unidades reducidas.
- (f) CELDAY,15.8 - Ancho en y de la celda original centrada en el origen en unidades reducidas.
- (g) CELDAZ,15.8 - Ancho en z de la celda original centrada en el origen en unidades reducidas.
- (h) CortaX,25.8 - Ancho en x de la celda centrada en el origen resultado de recortar el cluster en unidades reducidas. Sólo es necesaria si Flag_RecortarCluster=1. Dependiendo del tipo de periodicidad usada hay que comparar su magnitud con la de CELDAX para evitar problemas de periodicidad.
- (i) CortaY,25.8 - Ancho en y de la celda centrada en el origen resultado de recortar el cluster en unidades reducidas. Sólo es necesaria si Flag_RecortarCluster=1. Dependiendo del tipo de periodicidad usada hay que comparar su magnitud con la de CELDAY para evitar problemas de periodicidad.
- (j) CortaZ,25.8 - Ancho en z de la celda centrada en el origen resultado de recortar el cluster en unidades reducidas. Sólo es necesaria si Flag_RecortarCluster=1. Dependiendo del tipo de periodicidad usada hay que comparar su magnitud con la de CELDAZ para evitar problemas de periodicidad.
- (k) Flag_RecortarCluster,1 - Bandera que determina si se recorta o no el cluster. 0= No recortar el cluster. 1= Si recortar el cluste.

6. Referentes a la configuración de velocidades.

- (a) Tipo_Config_velocidades,0 - Opción para especificar el tipo de configuración de velocidades iniciales. Las velocidades se asignan con magnitud independiente de la temperatura inicial buscada y luego se escalan.

0= Todas las velocidades son 0, y la temperatura inicial es de $0^{\circ}K$. En este caso no es posible asignar una temperatura inicial. 1=La dirección y magnitud de las velocidades es asignada al azar. La semilla de la función de azar se toma del reloj interno de la máquina. 2= Distribución de velocidades inicial según una distribución de Maxwell. Es útil para sistemas muy inestables en los que los pocos pasos necesarios para desarrollar esta distribución pueden ser demasiados y destruir la configuración. 3=Velocidades radiales, proporcionales a la magnitud de la distancia del átomo al origen y con la misma dirección que un vector del origen al átomo. La proporcionalidad es la adecuada para generar la temperatura inicial requerida. Tienen complicaciones en cuanto a que aumentan los efectos de superficie para sistemas libres porque son precisamente los átomos más lejanos los que tienen mayor energía cinética. Sin embargo pueden ser útiles para ayudar a observar ciertas vibraciones, sobre todo porque mantienen cierta simetría y garantizan momento angular inicial cero.

- (b) TempInicial,2.0 - Parámetro para definir la temperatura inicial. Una vez asignadas las velocidades iniciales y ajustado el momento lineal total inicial a cero (en caso de que HacerMomentoInicialCero=1) las velocidades se escalan y se determinan su temperatura y energía total inicial. Este parámetro se usa para observar sistemas con la misma configuración atómica pero en diferentes estados (energías). Hay que recordar que las unidades no siempre son reducidas. Cuando un potencial es leído de archivo las energías se usan en eV y no en unidades reducidas.
- (c) Flag_Fijar_Algunos_Atomos,0 - Bandera que determina si todos los átomos van a ser considerados como móviles o si un subconjunto estarán restringidos a su posición inicial y energía cinética cero. 0=Todos los átomos son móviles. 1= Sólo algunos átomos son móviles (ie- hay átomos restringidos).
- (d) MovilX,0 - Ancho en x en unidades reducidas de la celda dentro de la cual se consideran los átomos como móviles. Debe ser menor o igual que CELDAX, de lo contrario es más eficiente hacer todos los átomos móviles.. No es necesario si Flag_SoloMedirAlgunosAtomos=0;
- (e) MovilY,0 - Ancho en y en unidades reducidas de la celda dentro de la cual se consideran los átomos como móviles. Debe ser menor o igual que CELDAY, de lo contrario es más eficiente hacer todos los átomos móviles.. No es necesario si Flag_SoloMedirAlgunosAtomos=0;
- (f) MovilZ,0 - Ancho en z en unidades reducidas de la celda dentro de la cual se consideran los átomos como móviles. Debe ser menor o igual que CELDAZ, de lo contrario es más eficiente hacer todos los átomos móviles.. No es necesario si Flag_SoloMedirAlgunosAtomos=0;
- (g) Flag_CenterInitialCluster,1 - Bandera que determina si se traslada el centro de masa del cluster inicial al origen. Esto se hace antes de asignar velocidades (importante para el caso de velocidades radiales). También es hecho antes de aplicar condiciones que dependen de las posiciones atómicas (como verificaciones de periodicidad, movilidad y otros). Es útil para ayudar a mantener constante el momento angular. 0=No se centra el cluster. 1=Si se centra el cluster.
- (h) Flag_HacerMomentoInicialCero,1 - Bandera que determina si el momento inicial se hace cero o no. Es útil para evitar derivas innecesarias y energías cinéticas adicionales que complican el análisis de los resultados. 0=No se hace cero el momento lineal inicial. 1=Si se hace cero el momento lineal inicial.

7. Referentes a las interacciones.

- (a) Tipo_Interacciones,0 - Opción para especificar el tipo de interacción, es decir, el potencial. 0=Lennard-Jones 12-6, 1=Morse, requiere definir Morse_Beta, 2=Pares con datos leídos de un archivo. 3= YYYe tipo átomo embebido, requiere definir ElementoEAM. 4=Gupta tipo átomo embebido, requiere definir ElementoEAM.
- (b) IfPossibleUseFormulas,1 - Bandera que determina que en caso de estar implementadas se usen las funciones analíticas en lugar de tablas para los cálculos de energía y fuerza. Sólo para los potenciales de pares leídos de archivo no hay funciones analíticas (por obvias razones). 0=Siempre usar tablas. 1=De existir usar funciones.
- (c) Min_R_Aceptable,0.7 - Parámetro relativo a la mínima distancia aceptable entre dos átomos para considerarlos superpuestos. En unidades reducidas. Es utilizado para evitar sobrecargas de la energía y en las verificaciones de periodicidad inicial.
- (d) Flag_Todos_Vecinos,0 - Bandera que determina si todos los átomos deben ser considerados vecinos ($R_{Out} \rightarrow \infty$). Hay que tener cuidado porque en sistemas grandes puede fácilmente desbordar el máximo de vecinos aceptables por átomo y hacer crecer los tiempos más allá de lo razonable. Recordemos que no sólo se incluye a cada átomo como vecino, sino que pueden serlo también sus 26 imágenes. 0= No todos los átomos son vecinos, es decir, listas de vecinos usuales. 1=Todos los átomos son vecinos.

- (e) RIn,6.5 - Radio de corte en unidades reducidas. Máxima distancia en unidades reducidas a la que se consideran las interacciones interatómicas.
- (f) ROut,8.5 - Radio exterior en unidades reducidas para listas de vecinos (listas de Verlet). Todos los átomos a una distancia menor o igual que ROut de un átomo *i* se consideran vecinos de este átomo. Es innecesario si Flag_Todos_Vecinos=1.
- (g) FactorActualizacion,4 - Parámetro para determinar la necesidad de actualizar listas de vecinos. Un valor mayor implica una actualización más seguido y mejor conservación de la energía. Lo inverso para menores valores. La actualización ocurre cuando $\max_{\text{átomos}}(x_t - x_{\text{última_actualización}}) \geq \frac{ROut - RIn}{\text{FactorActualización}}$.
- (h) Morse_Beta,1 - Parámetro necesario al utilizar el potencial de Morse.
- (i) ElementoEAM,Au - Para determinar los valores de las constantes al usar un potencial de tipo de átomo embebido.
- (j) SmoothCicles,5 - Número de veces que se itera la aplicación del algoritmo de suavizado del potencial de archivo. Más iteraciones permiten un mejor cálculo de las fuerzas, pero desplazan el valor mínimo y la distancia de equilibrio. Menos iteraciones pueden conducir en potenciales escalonados a calculos erroneos de las fuerzas. Sólo es necesario si el potencial es leído de un archivo.
- (k) SmoothPoints,5 - Número de puntos vecinos más cercanos usados en el algoritmo suavizado. El valor por defecto es 3. No tiene efecto si SmoothCicles=0.

8. Referentes a la salida del potencial.

- (a) Flag_Potencial,0 - Bandera que determina si se graba el archivo del potencial de pares utilizado. El archivo es de tipo xy y contiene en sus tres columnas los valores *r*, V(*r*) y F(*r*). 0=No se graba un archivo con el potencial. 1=Si se graba un archivo con el potencial.
- (b) Potencial_Out_Rmin,0.25 - Parámetro del valor mínimo de *r* (en unidades reducidas) a utilizar en las tablas internas y a grabar en el archivo de potencial. En combinación con Potencial_Out_Rmax y el número de puntos en las tablas determina la precisión con que se construyen. No es necesario que sea menor que Min_R_Aceptable.
- (c) Potencial_Out_Rmax,2.25 - Parámetro del valor máximo de *r* (en unidades reducidas) a utilizar en las tablas internas y a grabar en el archivo de potencial. Es necesario que sea igual o mayor que RIn.
- (d) Potencial_Out_Puntos,500 - Número de puntos a grabar en el archivo. De ellos se determinan los valores *r* que se graban. En el archivo no se graban todos los puntos de una tabla para que sean más fácilmente manejables por programas de graficación.
- (e) Potencial_Out_MaxOutput,5 - Parámetro con el valor absoluto de los valores máximos del potencial y la fuerza que se almacenan en el archivo de potencial (en las tablas internas el límite es el máximo de la máquina). Esto es para evitar sobrepasar los límites de programas de graficación y facilitar el aprovechamiento de la tabla generada.

9. Referentes a la periodicidad.

- (a) Tipo_Periodicidad,2 - Opción del tipo de condiciones de frontera. 0=Sistema libre.1=Fronteras de espejo. 2=Sistema periódico. Es importante recordar que si se utiliza un sistema libre los valores de Flag_Usar_Imagenes_X, Flag_Usar_Imagenes_Y y Flag_Usar_Imagenes_Z deben ser cero, por congruencia con la suposición de un sistema libre.
- (b) Flag_Usar_Imagenes_X,1 - Bandera que determina si se utiliza periodicidad en la dirección x. Esta bandera tiene efecto en diferentes lugares: i) Verificación de periodicidad inicial. ii) Condiciones de frontera en el movimiento y en las interacciones (en las listas de vecinos) y iii) Imágenes grabadas en los archivos de configuración (pdb's).
- (c) Flag_Usar_Imagenes_Y,1 - Bandera que determina si se utiliza periodicidad en la dirección y. Tiene efectos equivalentes a Flag_Usar_Imagenes_X.
- (d) Flag_Usar_Imagenes_Z,1 - Bandera que determina si se utiliza periodicidad en la dirección z. Tiene efectos equivalentes a Flag_Usar_Imagenes_X.
- (e) Flag_CuidarPeriodicidad,1 - Bandera que determina si se revisa la periodicidad inicial, retirando los átomos que no la cumplan. 0= No se hace ninguna verificación (cuidado, queda a discreción del usuario). 1) Si se hace verificación.

10. Referentes a las unidades.

- (a) `Real_Rmin_A, 3.816371` - Parámetro que permite determinar la correspondencia entre unidades reducidas y unidades reales. Es el valor de la distancia de interacción de mínima energía entre dos átomos. El programa supone que esta en angstroms. No es necesario si no se piden conversiones de unidades (valor por defecto).
- (b) `Real_Emin_Jmol, 997.68` - Parámetro que permite determinar la correspondencia entre unidades reducidas y unidades reales. Es el valor de la energía mínima de interacción entre dos átomos. El programa supone que esta en joules/mol. No es necesario si no se piden conversiones de unidades (valor por defecto).
- (c) `Real_Masa_kgmol, 0.0402897` - Parámetro que permite determinar la correspondencia entre unidades reducidas y unidades reales. Es el valor de la masa de un átomo. El programa supone que esta en kg/mol. No es necesario si no se piden conversiones de unidades (valor por defecto).

11. Referentes a las mediciones generales.

- (a) `Flag_SoloMedirAlgunosAtomos, 1` - Bandera que determina si se van a restringir algunos cálculos a los átomos medibles (un subconjunto de los átomos móviles). 0=Todos los átomos son medidos. 1=Sólo los átomos marcados como medibles son medidos.
- (b) `MediblesX, 3.0` - Ancho en x en unidades reducidas de la celda dentro de la cual se consideran los átomos como medibles. Debe ser menor o igual que `MovilX`. No es necesario si `Flag_SoloMedirAlgunosAtomos=0`;
- (c) `MediblesY, 3.0` - Ancho en y en unidades reducidas de la celda dentro de la cual se consideran los átomos como medibles. Debe ser menor o igual que `MovilY`. No es necesario si `Flag_SoloMedirAlgunosAtomos=0`;
- (d) `MediblesZ, 3.0` - Ancho en z en unidades reducidas de la celda dentro de la cual se consideran los átomos como medibles. Debe ser menor o igual que `MovilZ`. No es necesario si `Flag_SoloMedirAlgunosAtomos=0`;
- (e) `RawPeriod, 1` - Parámetro con el valor del período con el que se graban los valores "crudos" al archivo de salida. Sólo se decide si es momento de hacer una grabación de los valores cada `CICLOS_CIEGOS`, por lo que es necesario que para que el período de grabación coincida con lo aquí anotado que `RawPeriod` sea un múltiplo de `CICLOS_CIEGOS`.
- (f) `AveragesPeriod, 10` - Parámetro con el valor del período con el que se graban los valores "cocinados" al archivo de salida. Sólo se decide si es momento de hacer una grabación de los valores cada `CICLOS_CIEGOS`, por lo que es necesario que para que el período de grabación coincida con lo aquí anotado que `AveragesPeriod` sea un múltiplo de `CICLOS_CIEGOS`.

12. Referentes a las mediciones de distribuciones.

- (a) `Flag_DistPares, 0` - Opción para elegir el tipo de cálculo y la grabación del archivo de la distribución de pares. 0= No se hacen cálculos ni se graba archivo alguno. 1= Sólo se graba la distribución de pares inicial. 2=Sólo se graba la distribución de pares final. 3= Se acumula la distribución cada ciclo general y se graba la final promediada sobre el número de `CICLOS_GENERALES`. El archivo de salida es de tipo xy, con los valores de r y $N(r)$.
- (b) `Flag_DistVel, 0` - Idéntica que la opción arriba mencionada para la distribución de pares pero para la distribución de velocidades. El archivo de salida es de tipo xy, con los valores de r y $N(r)$.
- (c) `DistPares_Rmax, 5` - Máxima r (en unidades reducidas) acumulable en las distribuciones de pares. Usualmente es igual al radio de corte (R_{in}). En caso de que una distancia sea mayor que `DistPares_Rmax` se acumula el valor de la casilla de máxima r.
- (d) `DistPares_Rmin, 0` - Mínima r (en unidades reducidas) acumulable en las distribuciones de pares. Usualmente es igual a `Min_R_Aceptable`. En caso de que una distancia sea menor que `DistPares_Rmin` se acumula el valor de la casilla de mínima r.
- (e) `DistVel_Vmax, 10` - Máxima velocidad (en unidades reducidas) acumulable en las distribuciones de velocidades. En caso de que una velocidad sea mayor que `DistVel_Vmax` se acumula el valor de la casilla de máxima v.
- (f) `DistVel_Vmin, 0` - Mínima velocidad (en unidades reducidas) acumulable en las distribuciones de velocidades. En caso de que una velocidad sea menor que `DistVel_Vmin` se acumula el valor de la casilla de mínima v.

- (g) `DistPares_Num,100` - Número de casillas (diferentes intervalos de r) para las acumulaciones y grabado de las distribuciones de pares. Un valor muy alto tiene como resultado una distribución plana, con cada distancia en su clase particular. Un valor demasiado pequeño implica poca resolución en la curva y la imposibilidad de sacar mucha información de la medición.
- (h) `DistVel_Num,100` - Equivalente a `DistPares_Num`, pero para las distribuciones de pares.
13. El algoritmo de integración.
- (a) `Tipo_Integracion,1` - Opción para elegir el algoritmo de integración deseado. El más eficiente y mejor adaptado de los implementados es el de Verlet Velocidad, por lo que recomiendo usar 1. 0=Algoritmo de Verlet simple. 1=Algoritmo de Verlet Velocidad.
14. Respecto a cada una de las etapas. Los elementos siguientes se repiten para cada una de las 5 etapas posibles (etapa0 \rightarrow etapa4). Los que muestro son los de la etapa 0 y son sólo un ejemplo. Para el resto de las etapas se sustituye el 0 con el número de la etapa.
- (a) `Etapa0_Realizar_Etapa,0` - Bandera que determina si se realiza o no la etapa 0. 0=No se realiza y no son necesarios el resto de los parámetros. 1=Si se realiza y se toman como parámetros los valores que comienzan con `Etapa0`.
- (b) `Etapa0_CICLOS_CIEGOS,100` - Número de ciclos ciegos. Un ciclo ciego es la unidad mínima de avance en dinámica molecular. Durante un ciclo ciego el tiempo avanza DT . El número de ciclos totales para cada etapa es `CICLOS_CIEGOS*CICLOS_GENERALES`. Durante cada ciclo ciego se evalúan las fuerzas y energías, y se cambian las posiciones y velocidades de los átomos. Fuera de esto los únicos cálculos extra que se hacen son "el virial" para calcular la presión y que aprovecha los cálculos ya hechos en la rutina de las fuerzas, y la verificación de si es necesario actualizar o no las listas de vecinos. Un ciclo ciego está pensado para tener la máxima eficiencia posible. Una simulación de la que no requerimos mediciones podría consistir de puros ciclos ciegos. Por otro lado si requerimos mediciones constantes haremos `CICLOS_CIEGOS=1`. No hay que olvidar que las mediciones y grabaciones sólo pueden ser realizadas cada `CICLOS_CIEGOS`, por lo que hay que asegurarse que los períodos de estas sean múltiplos de `CICLOS_CIEGOS`.
- (c) `Etapa0_CICLOS_GENERALES,200` - Número de ciclos generales. Durante un ciclo general el sistema avanza `CICLOS_CIEGOS*DT` tiempo. Al final de los ciclos ciegos se hacen los cálculos del sistema, desde la medición de la energía cinética y el momento lineal, hasta las distribuciones de velocidades y las grabaciones de los cuadros de película.
- (d) `Etapa0_Tipo_Dinamica,1` - Opción del tipo de evolución que queremos que siga el sistema. El tipo de evolución se controla por medio de aplicar diferentes "ajustes". 0=Simulación clásica NVE, en la que no aplicamos ninguna clase de ajuste y el sistema evoluciona libremente. 1=Escalamiento fijo de velocidades. Cada `CICLOS_CIEGOS` las velocidades se escalan por un factor constante definido por `Escalador_cte_velocidades`. 2=Temperatura objetivo. Cada `CICLOS_CIEGOS` las velocidades se escalan por el factor adecuado para hacer que la temperatura del sistema sea la definida por `Temp_objetivo`. Esto sirve tanto para llevar al sistema a cierta temperatura como para mantenerlo fijo en una. La magnitud de `CICLOS_CIEGOS` y la temperatura objetivo determinan las fluctuaciones de temperatura entre escalamiento y escalamiento. A mayor temperatura y/o más `CICLOS_CIEGOS` mayores fluctuaciones.
- (e) `Etapa0_DT,0.001` - Intervalo de tiempo de integración. Es el tiempo transcurrido entre un ciclo ciego y el siguiente. Parámetro de las ecuaciones de movimiento que representa la evolución del tiempo. Depende del algoritmo de integración y del sistema, pero fijos estos por lo general se determina en función de las fluctuaciones en la conservación de la energía que estemos dispuestos a tolerar. DT está en unidades internas.
- (f) `Etapa0_Escalador_cte_velocidades,0.8` - Parámetro que representa el factor por el que se escalan las velocidades cada `CICLOS_CIEGOS` en caso de estar realizando una dinámica de escalamiento fijo de velocidades (`Etapa0_Tipo_Dinamica=1`). No es necesario si el tipo de dinámica es diferente.
- (g) `Etapa0_Temp_objetivo,0.02` - Parámetro que determina la temperatura objetivo cuando se está realizando una dinámica de temperatura objetivo. El sistema es escalado cada `CICLOS_CIEGOS` para que tenga esta temperatura. Es útil para llevar al sistema a alguna temperatura específica diferente de cero en forma rápida y para mantener la temperatura constante. Esta temperatura está en unidades internas. No es necesario si el tipo de dinámica es otro.

- (h) *Etapa0_Pdb_Out,E:\Molecular\PDBs_Out\Etpa0Out.pdb* - Nombre y dirección del archivo con formato pdb con la configuración al final de la etapa 0.

Cada uno de los elementos del archivo de instrucciones representa una forma de indicar al programa que realice o deje de realizar cierta función. Estas funciones pueden ser simples o muy complejas. No es posible dar aquí una explicación detallada del efecto de cada uno de los elementos, de la manera en que son considerados por el programa o de la forma en que deben elegirse. Incluyo aquí una descripción simple de lo que significan, una descripción detallada sería una iteración de esta tesis. Nombrar en cada elemento las secciones en donde se discute su propósito es también muy ambicioso y presenta algunos problemas: si incluyo todas las secciones en que se hace referencia a un elemento, para cada uno de ellos habría que listar 10 ó 15 secciones. Si me limito a mencionar las principales entonces no se gana mucho más que si el lector revisa el índice general de la tesis. Por lo tanto, en búsqueda de claridad no incluyo las referencias a las secciones.

9.8.2 Ejemplos de instrucciones

Cada circunstancia de interés involucra una combinación de instrucciones. A continuación doy algunos ejemplos de las combinaciones necesarias para casos conocidos. Las instrucciones que no menciono es porque no son función del caso en cuestión (es decir, no toman valores especiales ni deben estar sujetos a consideraciones particulares).

1. **Gradiente conjugado ($\bar{\nabla}$)**- Es un método de minimización de energía estable, pero poco eficiente. Para una descripción del método de gradiente conjugado ver la sección 6.11.3. Es un caso particular del enfriamiento suave cuando ocurre $\lim_{Escalador_cte_velocidades} \rightarrow 0$.
 - Tipo_Dinamica=1 (Escalar Fijo)- La energía cinética se eliminara completamente, y aunque podríamos hacerlo con Temp_objetivo=0 es menos eficiente porque serían hechos cálculos del factor de conversión.
 - CICLOS_CIEGOS - Debe ser muy pequeño. Usualmente no mayor de 1 o 2.
 - DT - Pequeño para evitar encontrar fuerzas grandes que den lugar a grandes velocidades (vg. 0.001)
 - CICLOS_GENERALES - Determina la duración de la simulación. Por lo general es relativamente grande (vg. algunos miles, pero depende del sistema), porque debido a la eliminación de las energías cinéticas el avance es muy pobre y son necesarias simulaciones largas.
 - Escalador_cte_velocidades=0.0 - Se eliminan por completo las energías cinéticas.
 - Temp_objetivo - Este parámetro no es tomado en cuenta por el tipo de dinámica seleccionada.
 - Otros - Los promedios pierden mucha de su utilidad y es recomendable dejar de realizarlos haciendo `AveragesPeriod>CICLOS_GENERALES*CICLOS_CIEGOS`.
2. **Enfriamiento suave** - Parecido al gradiente conjugado, pero menos drástico en la reducción de la energía cinética, buscando un balance entre la estabilidad del gradiente conjugado y la eficiencia en la búsqueda de configuraciones que permiten mayores velocidades.
 - Tipo_Dinamica=1 (Escalar Fijo)- La energía cinética se disminuye a una razón constante, lo que no implica una reducción constante en la energía total, puesto que la proporción K/E depende fuertemente del estado del sistema.
 - CICLOS_CIEGOS, DT y Escalador_cte_velocidades. Entre estos tres parámetros se encuentra el óptimo de eficiencia y estabilidad. Los valores adecuados dependen fuertemente del sistema. Por lo general CICLOS_CIEGOS~ 10 es adecuado con un DT=0.01. La velocidad del enfriamiento se controla con Escalador_cte_velocidades, el que normalmente no debe ser menor a 0.95 o el sistema llega rápidamente a comportarse como en un gradiente conjugado.
 - CICLOS_GENERALES - Se busca el adecuado para asegurar una convergencia a $0^{\circ}K$.
 - Temp_objetivo - Este parámetro no es tomado en cuenta por el tipo de dinámica seleccionada.
 - Otros - En este caso los promedios son particularmente importantes, si hacemos `AveragesPeriod = ciclos_ciegos` podemos averiguar la energía promedio en cada intervalo de energía constante, y usar esta información para elegir mejor la velocidad de enfriamiento. Entre mayor sea la energía cinética en un intervalo de CICLOS_CIEGOS, menos estable es el sistema y más eficiente la búsqueda del mínimo de energía.

3. **Temperatura constante** - Se elige una temperatura objetivo a la que queremos llevar al sistema, puede ser $0^{\circ}K$ para simular un gradiente conjugado, pero por eficiencia es más común utilizarla para temperaturas mayores. El método puede ser utilizado con dos diferentes objetivos en mente: i) Para llevar un sistema a una cierta temperatura objetivo (de allí al nombre del tipo de dinámica), calentándolo o enfriándolo según sea necesario. 2) Mantener una cierta temperatura. En cada uno de los casos se recomiendan parámetros diferentes.

(a) Para llevar al sistema a una temperatura objetivo

- Tipo_Dinamica=2 (Temperatura Objetivo)
- CICLOS_CIEGOS - Por lo general son mayores que en los casos gradiente conjugado y enfriamiento suave, para permitir al sistema cierto equilibrio antes de modificar su temperatura nuevamente. Si escogemos CICLOS_CIEGOS muy pequeño se está calentando demasiado rápidamente al sistema, lo que puede generar núcleos de fusión y modificar la configuración de forma importante. CICLOS_CIEGOS demasiado grande lleva a simulaciones poco eficientes.
- DT - Se elige en conjunción con CICLOS_CIEGOS para minimizar las fluctuaciones en la energía. Depende también de la temperatura objetivo. Una temperatura objetivo grande \Rightarrow energía cinética grande \Rightarrow necesidad de DT pequeño para mantener la validez de los métodos de integración. Por otro lado Temp_objetivo pequeña \Rightarrow energía cinética pequeña \Rightarrow posibilidad de usar DT grande y aumentar la eficiencia. Sin embargo el hecho de escalar las velocidades es ya una violación de la conservación de la energía, por lo que los efectos de un DT demasiado grande pueden no ser evidentes.
- CICLOS_GENERALES - Determina la duración de la simulación y acostumbra fijarse en base a la experimentación del tiempo requerido para que el sistema se equilibre a una temperatura.
- Escalador_cte_velocidades - Este parámetro no es tomado en cuenta por el tipo de dinámica seleccionada.
- Temp_objetivo - Este parámetro no es tomado en cuenta por el tipo de dinámica seleccionada.
- Otros - La medición de promedios puede ayudar a elegir adecuadamente el número de CICLOS_CIEGOS igual que en el enfriamiento suave.

(b) Para mantener constante la temperatura.

- Tipo_Dinamica=2 (Temperatura Objetivo)
- CICLOS_CIEGOS - Con este parámetro definimos el balance entre la eficiencia de la simulación y las fluctuaciones permitidas en la energía cinética. Un parámetro CICLOS_CIEGOS demasiado conlleva grandes fluctuaciones, pero una simulación más veloz. Un parámetro demasiado pequeño implica un excelente ajuste en la temperatura, pero a costa de tiempo de cálculo.
- DT - Se elige bajo las mismas consideraciones que cuando se lleva al sistema a una Temperatura objetivo.
- CICLOS_GENERALES - Si ya tenemos al sistema en la temperatura de interés usualmente es durante estos pasos que hacemos las mediciones definitivas del sistema, por lo que el número de CICLOS_GENERALES depende del tipo de observaciones que queramos hacer.
- Escalador_cte_velocidades - Este parámetro no es tomado en cuenta por el tipo de dinámica seleccionada.
- Temp_objetivo - Este parámetro no es tomado en cuenta por el tipo de dinámica seleccionada

4. **Celda periódica** - Es el método para simular sistemas infinitos descrito en la sección 6.3.

- Tipo_Periodicidad=1 o 2 (Espejo o periódica)- Lo más normal es escogerla periódica para evitar violaciones de la conservación de la energía y/o el momento por la interacción con paredes ficticias.
- Flag_Usar_Imagenes_X=1, Flag_Usar_Imagenes_Y=1, Flag_Usar_Imagenes_Z. Aún cuando no es necesario que se usen condiciones periódicas en las tres direcciones, de hecho para el caso de interés en esta tesis aplicado a fronteras de grano es lo que hicimos en la mayor parte de los casos, más no en todos.
- Si $CELDA_X \geq CortaX$, $CELDA_Y \geq CortaY$, $CELDA_Z \geq CortaZ$ y $Flag_RecortarCluster=1$ nos aseguramos de que no hay violaciones iniciales de las condiciones de periodicidad, lo que puede conducir a resultados equivocados o a una falla definitiva del programa.
- En caso de que no se cumpla lo anterior es importante hacer $Flag_CuidarPeriodicidad=1$, de lo contrario no se revisa al comienzo de la simulación que el cluster cumpla con las condiciones de frontera iniciales.

- `Imagenes_Grabar_Film >0` - No es indispensable que sea mayor que cero, pero es común para grabar en la película una muestra de al menos una celda imagen para poder observar algunos de los efectos de la condición de periodicidad.
 - Es recomendable que $RIn \leq \frac{1}{2} \sqrt{CELDAx^2 + CELDAY^2 + CELDAZ^2}$ para reducir los efectos de la periodicidad y la inducción de estructuras con cierta simetría.
5. **Celda restringida ("caja")** - Es un método utilizado en simulaciones de fronteras de grano reportadas en la literatura, y son precisamente sus efectos comparados con la Celda periódica los que queremos observar.
- `Tipo_Periodicidad=0` (Sistema libre) - Implica un cluster libre en el espacio.
 - `MovilX < CELDAX`, `MovilY < CELDAY` y `MovilZ < CELDAZ`. De esta forma se construye el *casarón* de átomos fijos sobre los que se concentran los efectos de superficie. Para activar este efecto es necesario hacer `Flag_Fijar_Algunos_Atomos=1`.
 - Habiendo generado una superficie de átomos con fijos, y desconociendo los efectos de esto sobre los átomos del resto del cluster, acostumbramos restringir los átomos sobre los que se hacen las mediciones a un subconjunto. Esto lo hacemos por medio de `MediblesX < MovilX`, `MediblesY < MovilY` y `MediblesZ < MovilZ`. En este caso el cambio se activa con la bandera `Flag_SoloMedirAlgunosAtomos=1`;

9.9 Formatos de entrada y salida

Para los casos en que es posible el programa usa formatos estándar para poder compartir los archivos con otras personas.

9.9.1 Configuración

Para el archivo de configuración elegimos el formato PDB. Este formato surge del *Protein Data Bank*, y por lo tanto ha sido optimizado para aplicaciones de bioquímica. En el apéndice B.1.1 describo cuidadosamente las características del estándar. Sin embargo ha sido utilizado desde hace ya varios años en el departamento de materia condensada y se tiene experiencia en trabajar con él. Además hay una gran cantidad de programas de análisis y visualización que pueden leer el formato.

Utilizamos una versión modificada localmente de PDB, en donde sustituimos propiedades que no necesitamos por otras que no están disponibles, por ejemplo, en lugar de utilizar los espacios reservados para el número de cadena a la que pertenece un átomo, almacenamos allí información al respecto de la movilidad que éste tendrá en la simulación. De esta forma podemos aprovechar las capacidades de programas como Rasmol para mostrar propiedades de archivos PDB para observar distribuciones en nuestro sistema 6.5.

Cómo la precisión en las posiciones almacenables en archivos PDB no es muy grande (apenas 4 decimales), cuando hubo necesidad de mayor detalle utilizamos el formato XYZ o películas.

Dado que la configuración es uno de los mejores resultados que podemos obtener de una frontera de grano, el programa permite almacenar en PDB las posiciones de las partículas en diferentes momentos. Además del PDB obtenido como configuración de entrada (en su caso), podemos observar los siguientes PDB's:

1. Inicial - Generado justo después de haber terminado de preparar el cluster. Permite observar el cluster definitivo a simular, con todas las modificaciones a su estructura, recortes, velocidades y demás. Esta también es útil para poder aprovechar los mecanismos para generar y modificar clusters sin hacer simulación, como un programa independiente.
2. Final de cada etapa - Al terminar cada una de las etapas se almacena un PDB. Esto permite dividir la simulación en etapas, salvo las complicaciones derivadas de la pérdida parcial de información de las velocidades.
3. Final general - Al final absoluto de la ejecución del programa se genera un PDB.

Como se puede ver en el diagrama de la figura 9-4 el cluster puede ser sujeto de muchas modificaciones una vez leído/generado, para garantizar ciertas propiedades necesarias durante la evolución. Algunas de estas modificaciones se hacen en banderas asignadas a cada uno de los átomos del cluster, por ejemplo la movilidad de un átomo puede ser cero o uno. Aprovechando que el formato PDB permite almacenar diferentes propiedades el programa graba en estos espacios algunas de las banderas específicas a nuestro programa.

En el archivo PDB se almacenan el número atómico, elemento y posiciones en las posiciones para ello definidas (que WebLab de MSI no respeta) y en posiciones no estándar el programa almacena (el número de columna supone que la primera columna es la #1):

- En la columna 26 un número 1 ó 2 que determina si el átomo grabado es de la celda original (1) o una imagen (2). En RasMol (o RasWin que es el mismo programa en su versión de la Universidad de California) podemos observar esto seleccionando la opción de coloración por grupo. En una instalación estándar los átomos de la celda original serán azules y los de las 26 celdas imagen serán rojos.
- En la columna 22 un número 1, 2 ó 3 que determina si el átomo grabado era inmóvil-no medible (0), móvil-no medible (1) o móvil-medible (2) (Notese que no hay átomos medibles-nó móviles). El patrón se repite en todas las celdas imagen.
- En las columnas 61-66 se almacena una fracción que representa la energía cinética del átomo. La coloración depende de las temperaturas de los demás átomos, siendo rojos los de máxima energía, azules los de mínima y verdes los intermedios. No es evidente el método que RasMol usa para elegir los colores, sin embargo no parece ser lineal así que sugiero no hacer inferencias cuantitativas de la representación.
- En caso de clusters generados por el programa todos los átomos son de oro.
- La numeración de los átomos esta pensada para garantizar que no hay dos átomos con el mismo número, aún cuando no sean los mínimos posibles ni esten ordenados en el PDB según el número. El orden en el PDB es tal que primero esta el átomo 1 (arbitrario), seguido de sus imágenes, luego el átomo 2 y sus imágenes, y así sucesivamente hasta el último átomo y sus imágenes.

Para una descripción del estándar del formato PDB ir a la sección B.1.1.

9.9.2 Películas

El formato de salida de películas depende del programa con que queramos observar la película. Los más comunes son ReView en Windows y AtomTV en Silicon Graphics, aunque WebLab y Cerius también tienen capacidad de animación. Dado que la mayor parte de la programación la hicimos en Windows dimos prioridad al formato de ReView, que describo en el apéndice B.2, sin embargo por la programación en objetos el cambio a otro formato no debe representar ningún problema.

9.9.3 Potenciales

El formato elegido para los potenciales (tanto de entrada como de salida) es el más simple posible, el XY. Consiste de dos columnas, en la primera el valor de x y en la segunda Y(x). Los programas de graficación y estadística que más utilizamos nosotros: fgraph y WPlot no tienen ningún problema. a diferencia de paquetes como Excel que son mucho más lentos y tienen límites insuficientes en el número de puntos graficables (fgraph también tiene un límite de alrededor de 8600 puntos. en WPlot nunca nos encontramos un límite). No requerimos que la separación entre los puntos de los potenciales sea homogénea.

Durante la ejecución el programa puede crear un archivo con el potencial utilizado (en caso de que sea de pares). esto permite no sólo verificar que se uso el adecuado, sino usar el programa de simulación como un transformador de potenciales, que puede aplicar interpolaciones y suavizados, además de generar tablas de funciones analíticas.

9.9.4 Mediciones

Los resultados de mediciones de propiedades del sistema también los almacena el programa en diferentes archivos. Estos archivos tienen un formato conocido como XY1Y2Y3..., consiste en una primera columna a la que corresponden los valores de x , seguida de varias columnas con los valores de diferentes funciones de x , $Y_1(x)$, $Y_2(x)$, etc. La ventaja de este formato es que facilita comparar tendencias en diferentes propiedades. por ejemplo observar que la energía cinética y potencial efectivamente se complementan sumando la energía total. Una desventaja del formato es que requiere que todas las funciones de x se evalúen en todas y las mismas x 's. Los programas no comprenden espacios vacíos en los archivos.

Egraph tiene un límite de 4 funciones de x , y no puede interpretar archivos que tengan más. WPlot tiene un límite de 32,000 funciones.

Lo que representa cada columna en los archivos de resultados varía de simulación a simulación, porque depende de las instrucciones específicas. Lo más común es que la coordenada x este representada por el tiempo y al menos una de las funciones Y sea la energía.

Valores crudos y cocinados.

El programa puede dar salida a lo que en el ini llamamos *valores crudos* y *valores cocinados*. Los valores crudos son mediciones directas instantáneas. Los valores cocinados son valores medios desde el valor de inicio. Los valores cocinados son muy útiles para estudiar propiedades que fluctúan, de las que queremos encontrar tendencias globales.

Para el archivo con las mediciones "crudas" la salida incluye las mediciones instantáneas del tiempo, la energía total, la energía cinética, la magnitud del momento lineal total y la presión en orden desde la primera hasta la 5ª columna. Las unidades de salida son reducidas o "internas" según lo explicado en la sección 9.7.

Para el archivo con mediciones "cocinadas" la salida incluye las mediciones de tiempo, energía total promedio y varianza en la energía total. La frecuencia con la que se actualiza esta información y el intervalo sobre el que se realizan los promedios se determinan mediante instrucciones del archivo ini.

Distribuciones

Las distribuciones de velocidad y distancia tienen un formato idéntico al del resto de las mediciones, excepto que en ellas el eje x lo ocupa la variable sobre la que se mide la distribución (velocidad o distancia radial respectivamente) y Y es la frecuencia absoluta.

9.9.5 Bitácora (Log)

A partir de que comienza la ejecución del programa se crea un archivo de tipo log, cuyo nombre se especifica en las instrucciones del ini y que contiene información de la simulación efectuada y respuestas del programa, así como posibles errores. En el log se escriben el archivo ini utilizado, las modificaciones a parámetros importantes, así como la configuración del cluster y un registro del éxito o fracaso de las diferentes etapas de la simulación. Es útil como registro de las simulaciones hechas y cómo ventana para encontrar errores en el programa o en las instrucciones.

El formato es libre, y sólo tiene sentido observarlo con editores de texto.

Parte. IV
Resultados

10. LA ECUACIÓN DE ESTADO UNIVERSAL PARA METALES

“Lo más incomprensible del mundo es que sea comprensible.” Albert Einstein.

En un fluido o gas la ecuación de estado relaciona la presión con el volumen y la temperatura del material. En el estado sólido los tensores de tensión y esfuerzo remplazan las variables termodinámicas de presión y volumen porque el material ya no es simétricamente esférico. Por ejemplo, para pequeñas tensiones elásticas, el tensor elástico constante relaciona los tensores de tensión y esfuerzo, dando lugar a la Ley de Hooke. La Ley de Hooke es entonces una ecuación de estado para pequeñas tensiones a una temperatura (o rango de temperaturas) en que las constantes elásticas no varían mucho. En el estado sólido también podemos preguntar ¿cual es la presión requerida para cambiar el volumen de un material en forma homogénea?. Nuevamente encontramos la presión \mathcal{P} como función del volumen y la temperatura T , es decir $\mathcal{P} = \mathcal{P}(V, T)$, lo que llamamos la *ecuación de estado del material*. Notese que en estado sólido es sólo una parte de una ecuación de estado completa.

Uno de los descubrimientos mas sorprendentes de los 80's fué que las ecuaciones de estado de muchos metales, cuando eran escaladas adecuadamente caían en una curva universal [Ros84] [Sut93]. Aún más sorprendente fué el descubrimiento de que la misma curva también describe la adhesión metal-metal y otros procesos en diferentes materiales, incluso no metálicos. Actualmente no se sabe que tan universal es la relación [Sut93]. De ser válida en los materiales y circunstancias que nos interesan calcular en esta tesis y con nuestro programa, podríamos considerar olvidarnos de utilizar diferentes potenciales, generalizando los resultados. Siendo obtenida de mediciones experimentales, la ecuación de estado debe ser reproducida por los diferentes potenciales a utilizar en las simulaciones. Cualquier potencial que no tuviera como efecto una ecuación de estado equivalente a la universal sería descartado como válido para simular, al menos en el rango de estados termodinámicos en que hubieran diferencias notables con la universal. De esta forma podríamos tener un método más de verificación de validez de potenciales. Estas son las razones por las que decidimos estudiar más a fondo el problema que llamamos de *potenciales generalizados*.

El mecanismo descrito a continuación es el descrito por Sutton en su libro [Sut93]:

Para comenzar definimos el *radio de Wigner-Seitz* r_{ws} , que es el radio de una esfera cuyo volumen es igual al volumen por átomo en el material. Si hay N átomos por unidad de volumen se cumple

$$N \frac{4\pi}{3} r_{ws}^3 = 1 \tag{10.1}$$

El radio de Wigner-Seitz se usa como escala de longitud para caacterizar la distancia interatómica en un metal.

La energía de amarre E_b es la diferencia en energía de un material con respecto a la energía del mismo número de átomos aislados. La energía de amarre tiende a cero cuando cuando los r_{ws} tiende a infinito (los átomos estan

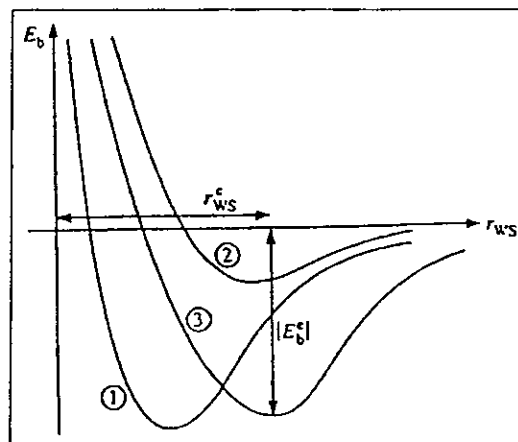


Figura 10-1 Energías de amarre E_b como función del radio de Wigner-Seitz r_{ws} para tres metales. El mínimo de cada curva ocurre en $r_{ws} = r_{ws}^e$ y $E_b = -|E_b^e|$. (cita al libro *Electronic structure of materials*).

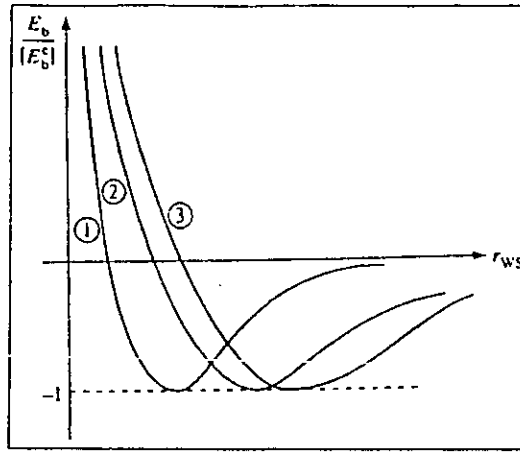


Figura 10-2 Energía de amarre escalada, $E_b/|E_b^e|$, como función del radio de Wigner-Seitz r_{ws} , para los mismos metales que la figura 10-1. Los mínimos ocurren en $E_b/|E_b^e| = -1$, pero para diferentes valores de r_{ws} .

infinitamente separados) y tiene un mínimo, negativo, en un valor de equilibrio $r_{ws} = r_{ws}^e$. La energía de amarre en el mínimo es E_b^e (ver figura 10-1). Diferentes materiales tienen diferentes curvas de energía de amarre porque tienen diferentes valores de equilibrio del radio de Wigner-Seitz y energía de amarre mínima. Lo sorprendente es que por medio de escalamientos de los ejes de longitud y energía se pueden colocar todas sobre una misma curva.

Primero escalamos el eje de la energía. Lo más simple es normalizar dividiendo entre el valor de la energía de amarre de equilibrio

$$g(r_{ws}) = \frac{E_b(r_{ws})}{|E_b^e|} \quad (10.2)$$

Las curvas $g(r_{ws})$ tienen mínimo en $(r_{ws}^e, -1)$ (fig. 10-2).

Para escalar el eje de longitud para que todas las curvas de la figura 10-2 estén una sobre otra, si simplemente restamos r_{ws}^e de r_{ws} y graficamos g como función de $r_{ws} - r_{ws}^e$ el mínimo sería llevado a $(0, -1)$, pero las curvaturas de $g(r_{ws} - r_{ws}^e)$ son diferentes y por lo tanto, las curvas no se traslapan, salvo en $(0, -1)$. Para tomar en cuenta este factor debemos considerar la segunda derivada de g en $r_{ws} = r_{ws}^e$. En la figura 10-2 se observa una aproximación parabólica en el mínimo. La longitud L se obtiene aproximando el mínimo por una parábola:

$$g(r_{ws}) \approx -1 + \frac{g''(r_{ws}^e)}{2}(r_{ws} - r_{ws}^e)^2 \quad (10.3)$$

que cruza el eje de longitudes cuando

$$r_{ws} - r_{ws}^e \approx \pm \sqrt{\frac{2}{g''(r_{ws}^e)}}. \quad (10.4)$$

Entonces,

$$L = 2\sqrt{\frac{2}{g''(r_{ws}^e)}} = 2\sqrt{2}\sqrt{\frac{|E_b^e|}{E_b''(r_{ws}^e)}} \quad (10.5)$$

El factor $2\sqrt{2}$ es irrelevante para definir un factor de escala y lo olvidamos. De esta forma escalamos $(r_{ws} - r_{ws}^e)$ por

$$l = \sqrt{\frac{|E_b^e|}{E_b''(r_{ws}^e)}} \quad (10.6)$$

y hemos definido una longitud escalada a como:

$$a = \frac{(r_{ws} - r_{ws}^e)}{l} \quad (10.7)$$

con la que podemos graficar g como función de a . El escalamiento de $E_b(r_{ws})$ resulta finalmente en el siguiente

$$E_b(r_{ws}) \rightarrow \frac{E_b\left(\frac{r_{ws} - r_{ws}^e}{l}\right)}{|E_b^e|} \quad (10.8)$$

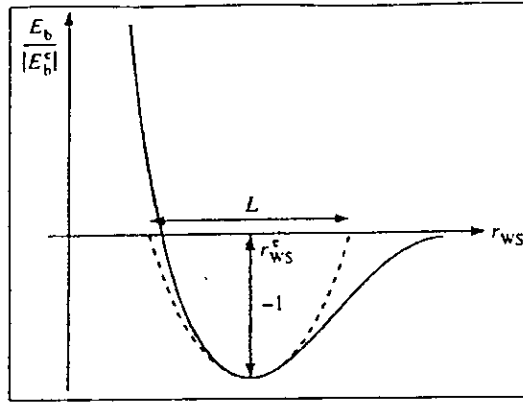


Figura 10-3 Aproximación por un potencial armónico.

	l (ec. 10.6)	r_{ws}^e	η
Oro	0.236	1.59	6.75
Plata	0.269	1.60	5.94
Platino	0.237	1.53	6.47

Tabla 10.1 Parámetros de escala para algunos metales [Ros84].

con lo que todas las curvas se traslapan, cómo se muestra en la figura 10-4.

Este resultado tiene conclusiones interesantes:

1. Conocidas la energía $|E_b^e|$, el módulo de bulto (*bulk modulus*) B y el radio de Wigner-Seitz para la densidad de equilibrio, se conoce la relación energía-densidad completa. El módulo de bulto en el equilibrio es

$$B = V \frac{d^2 E_b}{dV^2} = \frac{1}{12\pi r_{ws}^e} \frac{d^2 E_b}{dr_{ws}^2} \quad (10.9)$$

lo que implica

$$l = \sqrt{\frac{|E_b^e|}{E_b''(r_{ws}^e)}} = \sqrt{\frac{|E_b^e|}{12\pi r_{ws}^e B}} \quad (10.10)$$

2. La escala de longitud l es proporcional al ancho de la curva de energía de amarre, y por lo tanto el ancho de la región en que la aproximación parabólica es buena y la Ley de Hooke es válida. Un cristal es fuertemente anarmónico si esta región es pequeña. De esta forma podemos medir la anarmonicidad con

$$\eta = \frac{r_{ws}^e}{l} = \sqrt{\frac{12\pi B(r_{ws}^e)}{|E_b^e|}} \quad (10.11)$$

El texto [Sut93] incluye una tabla con los parámetros de escala l y de anarmonicidad η y el radio de Wigner-Seitz de equilibrio r_{ws}^e para diferentes metales. Algunos ejemplos son

Una vez encontrada la transformación que lleva a todas las curvas a traslaparse en una universal, podemos encontrar una forma funcional para la curva universal. La siguiente función es conocida como *función de Rydberg* y es una buena descripción [Sut93]

$$g(a) = -(1 + a + \frac{a^3}{20})e^{-a} \quad (10.12)$$

y todas las curvas en energía de amarre son descritas por

$$E_b(r_{ws}) = |E_b^e| g(a) \quad (10.13)$$

conocida como *Relación universal de energía de amarre* o UBER (por las siglas en inglés de *Universal Binding Energy Relation*). Algunos autores utilizan una versión simplificada de la ecuación 10.12, sin el término proporcional a a^3 .

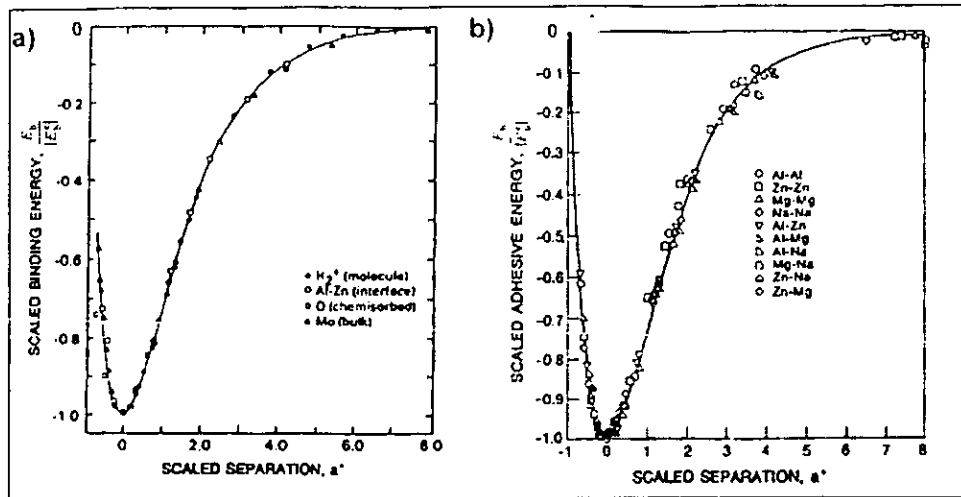


Figura 10-4 Energías de amarre escaladas vs radio de Wigner-Seitz. En ambas gráficas la línea continua es la función de Rydberg (ec. 10.12) comparada con valores experimentales. [Ban88]

10.1 Evidencia experimental

La universalidad de esta relación entre la energía de amarre y la separación interatómica ocurre para ligaduras covalentes y metálicas en muchas situaciones, desde moléculas diatómicas, adhesión bimetalica, chemisorption y cohesión en sólidos e incluso interacciones en materia nuclear [Ban88]. Las graficas de la figura 10-4 muestran una comparación entre la función de Rydberg y valores experimentales en diferentes situaciones.

10.2 Nos preguntamos

Con una cierta relación universal entre la energía de amarre y la separación entre partículas podríamos simplificar notablemente el análisis de simulaciones atómicas por dinámica molecular, pero sobre todo de los efectos de diferentes potenciales en las configuraciones de mínima energía de fronteras de grano generadas por el modelo GCSN. Para conocer si es posible aplicar esta relación a las simulaciones relacionadas con esta tesis necesitamos responder:

1. ¿Que tan general es la relación?. Esta pregunta es en el más amplio de los sentidos. Si la relación es completamente universal, es indispensable conocerla porque representa un avance fundamental en el estudio de potenciales interatómicos. Si no lo es, y se refiere solamente a ciertos materiales y/o bajo ciertas circunstancias, entonces cobra importancia conocer los límites de validez.
2. ¿Es posible aprovechar una relación general y construir de ella un potencial aprovechable en simulaciones?. Dado que cualquier curva de energía de amarre resultado de la aplicación de un potencial puede ser transformada a una curva supuestamente universal, quisieramos saber si existe una transformación inversa que, dados ciertos parámetros, se pueda construir un potencial interatómico para simular, o si es posible inclusive simular utilizando directamente la relación universal en lugar de un potencial. Acercamientos como estos permitirían una mayor generalidad en los resultados y conclusiones obtenibles de simulaciones.
3. ¿Que características debe tener un potencial interatómico para poder dar lugar a una curva transformable en la curva universal?. Respondiendo a esta pregunta podemos encontrar requisitos que deben satisfacer potenciales que pretendan modelar correctamente los materiales que se sabe cumplen esta relación. Así podríamos haber encontrado un método para distinguir las regiones en que los potenciales son útiles.

10.3 El procedimiento (algoritmo) para estudiarlo

Decidimos responder a las anteriores preguntas haciendo cálculos numéricos para encontrar las relaciones entre curvas de energía de amarre vs radio de Wigner-Seitz y la función de Rydberg (relación universal de energía de amarre).

Un programa que haga este tipo de cálculos consta de las siguientes partes:

1. Preparar el potencial y sistema atómico sobre el que se harán las mediciones. En esta parte generamos clusters cristalinos FCC con un átomo en el origen de coordenadas considerado como el átomo de prueba. En las mediciones supondremos sistemas infinitos, por lo que la energía de amarre total es $E_{total} = E_{un\ átomo} * Número\ de\ átomos$ y podemos calcularla calculando la energía de un átomo. Al aplicar el escalamiento en que se divide por la energía de equilibrio (ec. 10.2) cancelamos la aproximación y el resultado obtenido es exacto e idéntico al que se hubiera obtenido calculando la energía de amarre de todos los átomos. La preparación del potencial significa preparar las funciones analíticas que lo representan, leer los valores del potencial de un archivo en disco y/o crear y llenar las tablas a utilizar para ahorrar tiempo.
2. Encontrar la curva de energías de amarre como función del radio de Wigner-Seitz. Recorremos los radios de Wigner-Seitz en forma homogénea, encontramos para el valor de r_{ws} y la estructura atómica en consideración la escala atómica adecuada. Escalamos el cluster atómico por este valor y hacemos un cálculo de la energía de amarre del átomo del origen. Esto significa que no podemos hacer cálculos de este tipo para cualquier tipo de potencial, pero es útil para potenciales de pares y de tipo átomo embebido. Para calcular las distancias necesarias para el cálculo de la energía utilizamos una tabla de distribución de pares asociada a la estructura cristalina FCC. De esta forma no necesitamos escalar las distancias entre pares de átomos para cada par, sino que sólo escalamos la distancia en la tabla. La energía obtenida de esta distancia la multiplicamos por el número de átomos a esa distancia. Para potenciales de átomo embebido aprovechamos también tablas de distribución de pares, pero en una forma más complicada. Finalmente agregamos la posibilidad de usar estructuras BCC y cúbica simple.
3. Escalar la curva encontrada, transformarla y compararla con la función de Rydberg. Encontrada la relación E_b vs r_{ws} encontramos $E_b^e = E_b(r_{ws}^e)$, r_{ws}^e y $E_b''(r_{ws}^e)$. Utilizamos el procedimiento discutido atrás y transformamos la curva en una versión comparable con la universal. El punto mínimo y la segunda derivada los encontramos por métodos numéricos estándar. Para la segunda derivada utilizamos el método simétrico de 5 puntos descrito en la sección 8.5.2. Allí muestro un resultado interesante que encontramos al aplicar el método para calcular la curvatura en este lugar: encontramos que no era conveniente utilizar la separación entre puntos como la h mínima para calcular las derivadas, sino que si utilizamos valores entre $2h$ y $4h$ podemos encontrar valores mejores. Cuando usamos h mínima encontramos segundas derivadas que varían en forma discontinua de forma importante. La figura 8-6 muestra una comparación en el cálculo de la segunda derivada para el mínimo de un cristal FCC interactuando por un potencial de Lennard-Jones. Sin esta adaptación los valores de curvatura no son adecuados para hacer la transformación a la curva ideal.

Al igual que el programa de dinámica molecular este fue hecho orientado a objetos, de tal manera que las clases creadas pudieran ser compartidas, ahorrando tiempo y mejorando la claridad.

La manera en que el programa recibe instrucciones por medio de un archivo ini y los formatos de entrada y salida son idénticos a los del programa de dinámica molecular, convirtiendo a este programa en una versión muy reducida de aquel, útil para diversas pruebas de eficiencia y orden. También programe en Visual Basic una interface visual para Windows 95/98/NT4 que facilita el uso y aclara las diferentes posibilidades.

10.4 Resultados

Usando el programa especial calculamos las diferentes curvas de energía reducidas para los potenciales de pares principales de uso en problemas de fronteras de grano. La gráfica a de la figura 10-5 muestra los resultados para los potenciales de Lennard-Jones (sec. 7.4.3) y Morse con $\beta = 1$ (sec. 7.4.4). En la misma gráfica usamos un potencial de Rydberg (no confundir con la función ideal de Rydberg!) para aluminio del que no conocemos la forma analítica, pero que conocemos como tabla. Los tres potenciales concuerdan muy bien con la función de Rydberg

(ideal), aunque aparentemente el de Rydberg se aleja un poco más. Esta desviación esta dentro de los parámetros aceptables si observamos los experimentos que respaldan la universalidad de la curva (fig. 10-4). La gráfica b de la misma figura (10-5) muestra los resultados para potenciales de tipo de átomo embebido y Gupta (sección 7.6.1). Nuevamente las concordancias con la curva ideal son excelentes.

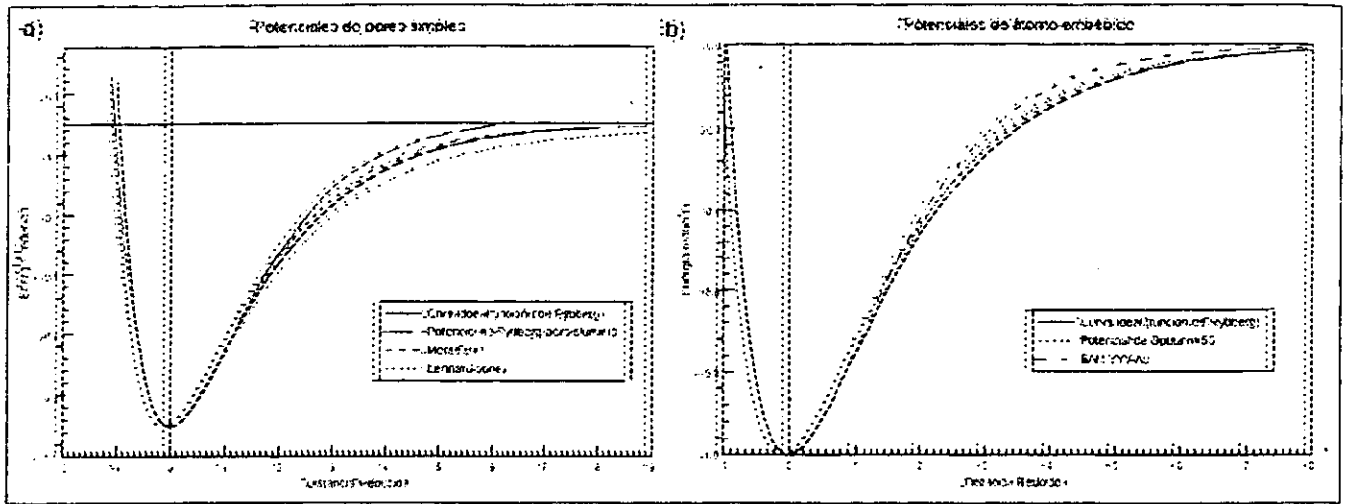


Figura 10-5 Curvas de energía de amarre calculadas e ideal para a) Potenciales de pares sin joroba. b) Potenciales tipo embeded atom (EAM). Se usaron configuraciones FCC.

Observando los resultados de la figura 10-5 podemos pensar que cualquier potencial va a tener una buena correspondencia con la curva ideal. Por ello decidimos continuar con las pruebas utilizando potenciales de pares con una diferencia importante: una joroba, correspondiente con una región de repulsión más allá del punto de equilibrio, y fuerzas repulsivas de allí hasta $r \rightarrow \infty$. En el máximo de la joroba hay un punto de equilibrio inestable. La gráfica a de la figura 10-8 muestra como la curva de energía de amarre reducida, contra distancia reducida mantiene esta forma. De esta forma el grupo de potenciales con la forma del de Bristowe para cobre mostraron NO corresponder con la curva supuestamente ideal.

Habiendo encontrado potenciales que no cumplen con la curva ideal, pero que tienen una forma esencialmente distinta, decidimos averiguar si potenciales de pares sin joroba, pero que no sean conocidos pueden dar lugar a curvas equivalentes a la ideal. Para esto utilizamos un potencial triangular como el de la figura 10-6. Las derivadas de este potencial no son contínuas y esto puede representar problemas, por lo que aplicamos el método de suavizado con 100 iteraciones 8.5.1. Encontramos que el potencial triangular tiene una buena correspondencia con la curva ideal.

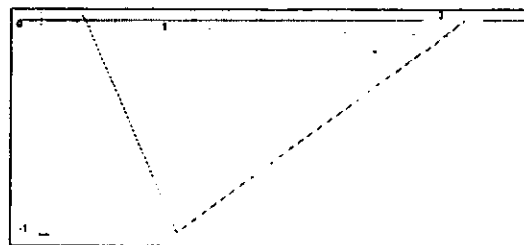


Figura 10-6 Potencial triangular usado para observar los potenciales que pueden coincidir con la función de Rydberg.

La misma gráfica b de la figura 10-8 muestra el resultado de aplicar un potencial de pares con la forma de la figura 10-7. La curva es el contorno de una montaña del estado norteamericano de Wyoming con la vertical invertida. La concordancia con la curva ideal no es tan buena como el triangular, pero respetable. Probablemente existan montañas con perfiles de mejor correspondencia. Nuevamente aplicamos un algoritmo de suavizado a la curva digitalizada del perfil de la montaña, pero fueron necesarias 1000 iteraciones para obtener la correspondencia adecuada.

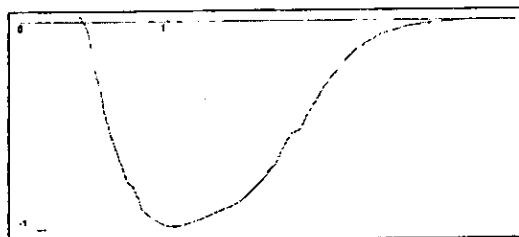


Figura 10-7 Potencial triangular usado para observar los potenciales que pueden coincidir con la función de Rydberg.

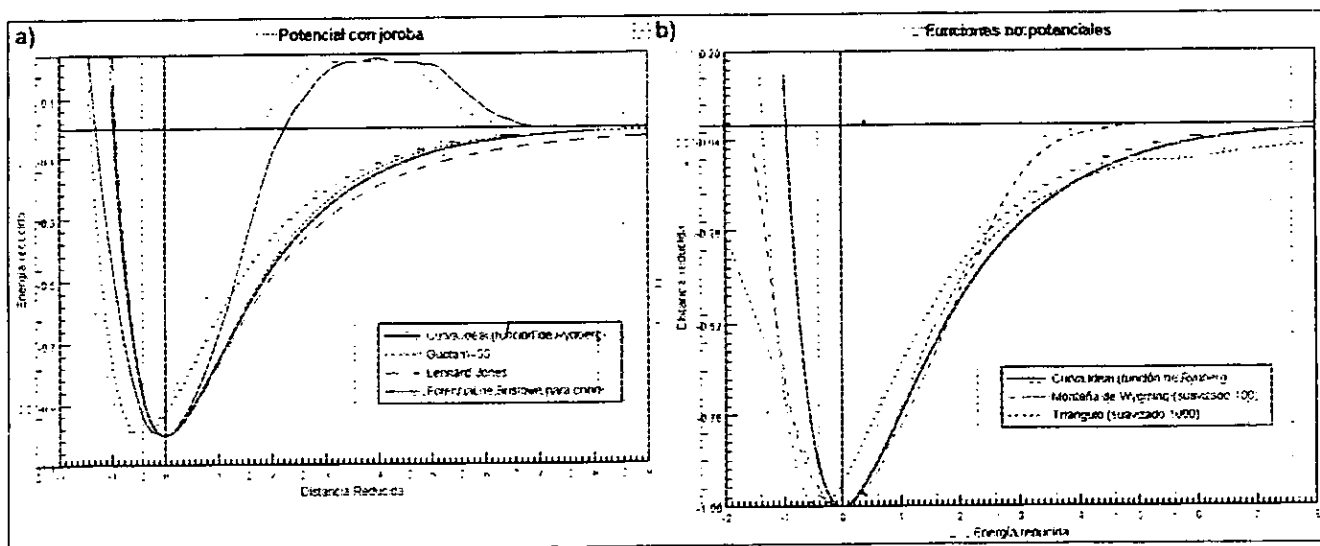


Figura 10-8 Energías de amarre de potenciales a) de pares con joroba y b) de otras figuras no potenciales consideradas como de pares. Notese que en el caso del potencial de Bristowe para cobre la forma general se preserva al calcular la energía de amarre. Las figuras usadas en b fueron sometidas a un suavizado para eliminar discontinuidades en las derivadas. Se usaron configuraciones FCC.

10.5 Conclusiones

Después de describir el mecanismo de transformación descrito en ROS84 para transformar curvas de energía de amarre en una curva universal (función de Rydberg) y aplicarlo a cálculos con diferentes potenciales interatómicos en estructuras FCC encontramos las siguientes virtudes y defectos de esta teoría:

Virtudes:

- Es adecuada cuando se calcula la energía de amarre de materiales por medio de expansiones homogéneas, en las que no hay esfuerzos de corte.
- Se llega a ella mediante un mecanismo fácil de aplicar, lo que facilita su aprovechamiento.
- Reproduce muchos materiales de diferentes tipos, desde metales a moléculas diatómicas.
- Permite conocer la curva de energía de amarre de un material conociendo algunos parámetros medibles experimentalmente.

Defectos

- Curvas que no representan potenciales realistas pueden dar lugar a energías de amarre que también corresponden con la universal.
- Potenciales que han demostrado hacer predicciones exitosas de algunas propiedades, como el de Bristowe para aluminio, no dan lugar a curvas transformables en la universal. Una joroba en el potencial lleva a una joroba en la curva de energía.
- Cualquier curva con un mínimo suave puede ser escalada para aproximarse cerca del mínimo a la función de Rydberg.

De todo esto concluimos que no es posible utilizar la función de Rydberg como sustituto para utilizar potenciales en una simulación de dinámica molecular, por los esfuerzos cortantes presentes en las interacciones. Sin embargo, es una relación interesante entre diferentes propiedades de los materiales. Permite utilizar algunos valores conocidos de materiales para conocer otros.

La teoría ayuda a entender algunos resultados de estado sólido. Banerjee y Smith discuten [Ban88] los orígenes de la relación para tan diferentes situaciones, considerando la densidad electrónica como la variable que relaciona la energía total y la separación entre partículas.

11. FRONTERAS DE GRANO

De acuerdo con lo expuesto en el capítulo 2.3 esperamos que las fronteras GCSN sean de baja energía comparadas con las fronteras de otros modelos, es el momento de utilizar el programa para estudiarlas. Lo primero es estudiar algunas de las propiedades simples en forma cualitativa. Simples en el sentido de que pueden ser observadas directamente (por medio de una película o animación) o medidas con parámetros estáticos. Comenzamos por estudiar propiedades que puedan ser comparadas con las conocidas experimentalmente y con las calculadas (y calculables) según otros modelos de frontera de grano. Por lo anterior decidimos estudiar la estabilidad de las fronteras GCSN.

Hay muchas maneras de estudiar la estabilidad porque hay diferentes aspectos que pueden representarla. Como primer cálculo de fronteras de grano con el programa de dinámica molecular generado mediante esta tesis hicimos enfriamientos simulados que llevaran las fronteras a una configuración de equilibrio de mínima energía potencial local a $0^\circ K$.

Después medimos en forma aproximada la temperatura de recristalización de una de las fronteras, para poder compararla con la experimental.

Finalmente observamos la aparición de algunas simetrías interesantes que pudieran representar modos normales de vibración y estudiamos esta posibilidad.

11.1 Los experimentos

11.1.1 Nuestras intenciones

Cuando las diferencias entre las posiciones de los átomos de una configuración con respecto a los de otra son mucho (varios órdenes de magnitud) menores que las distancias interatómicas decimos que las configuraciones son parecidas. Si mediante los enfriamientos llegamos a configuraciones parecidas a las iniciales, mostramos que las fronteras GCSN son buenos puntos de inicio (en el espacio de configuraciones) en cuanto a energía se refiere.

El algoritmo de enfriamiento por factor constante es sensible a fallas en la integración de las velocidades. Estas pueden aparecer en los experimentos con fronteras GCSN por las fuerzas rápidamente variables entre los átomos distancias pequeñas (menores que 1 en unidades reducidas, como las de la figura 11-21). Sin embargo de entre los algoritmos descritos en la sección 6.11 decidimos utilizarlo porque permite un muestreo del espacio fase más amplio que los algoritmos que no integran las velocidades, como el de gradiente conjugado. El enfriamiento por factor constante consiste en avanzar en el tiempo mediante una simulación clásica NVE, reduciendo cada cierto número de pasos la energía cinética por un factor constante menor que uno (ver sección 6.12.1).

Una prueba de estabilidad adicional es mantener el sistema a una temperatura constante (sección 6.14.2). A temperaturas pequeñas (por ejemplo 1% de la temperatura de fusión) una frontera GCSN debe ser estable incluso durante simulaciones de larga duración (por ejemplo 10 veces el tiempo en que llega al equilibrio (sección 6.4)).

Se sabe de los experimentos que la temperatura de fusión de un material con frontera de grano es menor que la de un material cristalino (sección 1.4). Para observar este fenómeno simulamos la evolución de la frontera a temperatura constante durante un tiempo al menos 10 veces mayor al de equilibrio. Lo anterior lo repetimos para diferentes temperaturas, desde $0^\circ K$ hasta temperaturas mayores que la de fusión del cristal. Esperamos observar que la estructura de la frontera se preserva para temperaturas ambiente (tal y como se observa experimentalmente (sección 1.4)), pero también que hay límites de temperatura máxima para la estabilidad de las fronteras.

Una frontera de grano se pierde a temperaturas menores que la de fusión. La temperatura a la cual que esto ocurre se conoce como *temperatura de recristalización*. El procedimiento para generar fronteras GCSN elimina algunos átomos del plano de la frontera, reduciendo la densidad y eliminando la posibilidad de que en las simulaciones a volumen constante podamos observar una recristalización completa. Sin embargo si es posible observar la pérdida de identidad de la frontera, de los motivos y de sus simetrías. Por lo tanto no pretendemos medir con precisión la temperatura de recristalización, sino mostrar que para los materiales simulados se encuentra entre la temperatura ambiente y la de fusión del material cristalino, tal y como se observa experimentalmente.

En resumen nos interesa saber:

- De las fronteras de grano:
 - ¿Son las fronteras GCSN estables a $0^{\circ}K$?
 - ¿Tienen las fronteras GCSN configuraciones cercanas a una de equilibrio?
 - ¿Son las fronteras GCSN estables a temperaturas ambiente?
 - ¿Es la temperatura de recristalización de las fronteras del orden de magnitud de la experimental?
 - ¿Hay diferencias en la estabilidad de las fronteras GCSN y las *de la competencia*?
 - ¿Que significa la aparición de simetrías nuevas? (ver sección 11.2)
- De los métodos de simulación de fronteras de grano:
 - ¿Hay diferencias notables/importantes en usar potenciales de pares *con joroba*, *sin joroba* o de tipo EAM?
 - ¿Es necesario utilizar potenciales de tipo EAM?
 - ¿Hay efectos notables del uso de condiciones de frontera periódicas?
 - ¿Cuales son los parámetros de enfriamiento y cálculo de interacciones óptimos?
- Del programa:
 - ¿Funciona adecuadamente?
 - ¿Es eficiente?
 - ¿Es fácil de usar?

11.1.2 Fronteras de interés

Para poder verificar las leyes de conservación y observar el funcionamiento del programa elegimos usar condiciones de frontera periódicas. Esto significa que debemos limitarnos a estudiar fronteras de grano con periodicidad pequeña. Pequeña en el sentido de que la cantidad de átomos contenidos en 8 planos más el de la frontera para unos cuantos períodos (digamos 3×3 ó 5×5) sea de unos cuantos miles, cantidad fácilmente tratable por nuestro programa.

Lo anterior lo satisfacen $\Sigma 5$ y $\Sigma 13$. Estas dos fronteras tienen además la ventaja de ser ampliamente conocidas en la literatura de fronteras de grano [Cha61] y de presentar motivos claramente identificables. Los motivos permiten observar que la periodicidad sea reproducida en forma correcta por el programa: se deben observar exactamente los mismos resultados para cada uno de los períodos.

Elegidas las fronteras a utilizar decidimos hacer una variante más que nos permitiera entender mejor lo que estamos observando: repetir todos los experimentos para dos diferentes estados traslacionales. Los estados traslacionales que utilizamos son los llamamos t_{csl} y t_2 , según la nomenclatura que se les ha asociado desde que Bollman [Bol74] les diera tratamiento matemático.

En resumen, como punto de inicio utilizamos las fronteras $\Sigma 5 t_2$, $\Sigma 5 t_{csl}$, $\Sigma 13 t_2$ y $\Sigma 13 t_{csl}$ generadas según el modelo GCSN.

Esperamos observar que $\Sigma 13$ tiene menores energías que $\Sigma 5$, tal y como se observa experimentalmente y encuentran otros métodos de simulación. Por lo general las traslaciones t_{csl} tienen menor energía que las t_2 , sin embargo puede que no siempre ocurra así, según explica Bollman.

Para fronteras de $\Sigma 5$ utilizamos alrededor de 1600 átomos (dependiendo de la traslación específica), mientras que para fronteras $\Sigma 13$ usamos alrededor de 1050 átomos.

11.1.3 Método de enfriado

Como explico en la sección 6.11 el método de enfriamiento por factor constante consiste en multiplicar las velocidades por un factor constante menor que uno cada cierto número de pasos. Esto no garantiza un enfriamiento por cambios de temperatura constantes, pero con el tiempo hace converger al sistema a una temperatura de $0^{\circ}K$. Realiza un muestreo del espacio fase porque permite movimientos de los átomos a posiciones lejanas, dependiendo de las fuerzas

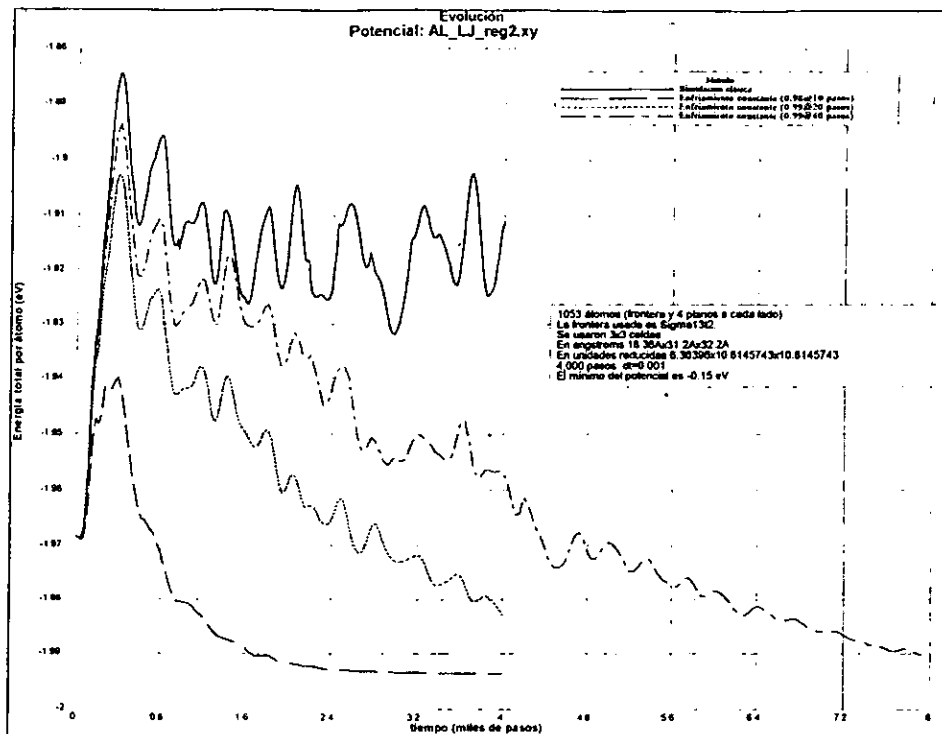


Figura 11-1 Pruebas de diferentes velocidades de enfriamiento de la frontera $\Sigma 13t2$ y el potencial de Lennard-Jones para aluminio. Con simulaciones como estas determinamos los parámetros adecuados para las minimizaciones posteriores. También muestran una evolución de energía constante para poder comparar. Notese que la energía mínima es muy cercana a la de inicio.

sobre ellos. Sabemos que es más lento que un método de gradiente conjugado, y ligeramente más complicado, pero para encontrar buenos mínimos locales vale la pena utilizarlo.

El método de control de temperatura por escalamiento con factor constante requiere de dos parámetros. Uno es el factor por el que se van a escalar las velocidades y el otro la frecuencia del escalamiento. Para una discusión de la forma en que deben elegirse los factores (escala de velocidades y frecuencia de aplicación) ver la sección 6.12. Para elegir los factores adecuados se hicieron varios experimentos, algunos de los cuales se muestran en la figura 11-1. La combinación de escalar por 0.98 cada 10 pasos de integración con $dt = 0.001$ resultó la que llevo a las configuraciones de menor energía. Si en el intervalo entre escalamientos la energía cinética aumenta es que el sistema se encuentra en una región de crecimiento de energía cinética. Conforme el sistema se acerca al equilibrio la ganancia de energía cinética disminuye.

La temperatura inicial de las simulaciones es $0^\circ K$, igual que la temperatura final. Esto complica determinar una *velocidad de enfriamiento*. Sin embargo podemos graficar tanto la temperatura como las variaciones en la temperatura (enfriamiento). Mediante el estudio del enfriamiento podemos observar la forma en que se comporta el sistema. El período de las oscilaciones en la temperatura puede darnos una indicación de un tiempo intrínseco del sistema para poder saber mejor si la duración de la simulación y la velocidad de enfriamiento son adecuadas.

11.1.4 Método de temperatura constante

En los experimentos en que se mantuvo constante la temperatura el método elegido fué el de esclamiento directo, en el que el factor de escala se calcula para cada actualización de manera que lleve a la temperatura objetivo. El método lo describo en la sección 6.12. Tiene la ventaja de ser simple y claro, pero requiere un período de equilibrio en el cual las mediciones no son muy confiables. La duración de este período aumenta al aumentar la temperatura objetivo. En todos las simulaciones se verifico que la duración siempre fuera al menos 10 veces mayor que el período de equilibrio.

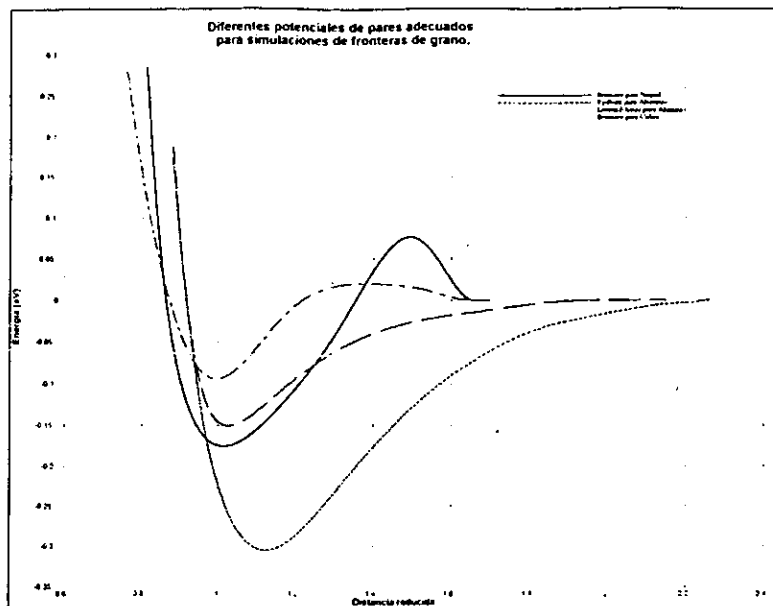


Figura 11-2 Cuatro potenciales de pares adecuados para la simulación de fronteras de grano en distintos materiales. Notese que se pueden distinguir claramente los dos que presentan *joroba*, de los dos que no la presentan.

No debemos olvidar que un algoritmo como el elegido implica una violación de la conservación de la energía, lo que puede contribuir a “escalar” pozos de potencial y destruir una configuración.

11.1.5 Potenciales

En el capítulo 7 explico como el potencial de interacción elegido determina, en gran medida, el grado de correlación con la realidad de una simulación. Sabemos que los potenciales que utilizamos son aproximaciones y queremos elegir aquel que sea lo más simple y que a la vez lleve a resultados que coincidan con la realidad, cualquier cosa que esto signifique.

Para tratar problemas de fronteras de grano han sido desarrollados varios potenciales de pares. Estos no sólo son particularmente simples, sino que han mostrado reproducir aceptablemente (al menos en forma cualitativa) resultados experimentales [Rom99] [Vot96]. Seleccionamos cuatro de estos y los mostramos en la figura 11-2. Los cuatro pueden ser divididos en dos grupos: con *joroba* y sin *joroba*. La *joroba* es la región de equilibrio inestable más allá de la de equilibrio estable. Esta región hace del punto de mínima energía un punto “más atractivo” para los átomos vecinos, aumentando notablemente la energía de los que se alejan. A mayores distancias que la de equilibrio inestable los potenciales tienen una región repulsiva mientras convergen lentamente a cero. Esto significa que la región de atracción es limitada, y todos los átomos fuera de ella son repelidos, a diferencia de los potenciales de pares sin *joroba*, donde la región de atracción se extiende hasta infinito.

Los parámetros de los cuatro potenciales han sido ajustados para representar lo mejor posible las propiedades de fronteras de grano. Los potenciales *sin joroba* son potenciales de pares simples de familias ampliamente conocidas en la simulación.

Para los experimentos con las fronteras $\Sigma 5$ y $\Sigma 13$ en sus dos estados traslacionales elegimos un potencial de cada grupo. El potencial de Bristowe para níquel *con joroba*, y el potencial de Rydberg para aluminio *sin joroba*. Los elegimos porque son los que tienen más marcada la forma: dentro de los *jorobados* el de Níquel tiene un máximo y mínimo de mayor magnitud que el de cobre, al igual que el de Rydberg tiene un mínimo con mayor magnitud que el de Lenard-Jones. Habría ventajas en utilizar otras combinaciones e hicimos también algunos cálculos con los otros potenciales, pero los resultados no mostraron diferencias notables.

Los cuatro potenciales los conocemos únicamente mediante un archivo con valores $V(r)$. La forma de utilizar un potencial de archivo la discuto en la sección 8.5.1. Hay una dificultad especial al aplicarlos a fronteras GCSN, y es que estas fronteras presentan parejas de átomos a distancias pequeñas, inclusive menores a las mínimas en los

archivos de potencial. Tuvimos que implementar un método de extrapolación para distancias pequeñas. Para evitar que aparecieran fuerzas muy grandes que pudieran hacer fallar el algoritmo de integración a distancias pequeñas las fuerzas son iguales a la de mínima distancia, y las energías iguales también a la de mínima distancia. Esto no es correcto físicamente y repercute en una subestimación de la energía, sin embargo diferentes cálculos mostraron que no tiene efectos notables en el resultado final.

En años recientes ha surgido un renovado interés por los potenciales de muchos cuerpos, en particular por los del grupo de *átomo embebido* (sección 7.6.1). Este interés ha sido propiciado por diferentes factores: 1) Se ha mostrado que hay propiedades que no pueden ser correctamente predichas mediante potenciales de pares. 2) Nuevas computadoras, mucho más poderosas que abren la posibilidad de realizar los cálculos extra requeridos por estos potenciales. 3) Mayor interés en predecir de forma cuantitativa propiedades que anteriormente bastaba con predecir de forma cualitativa.

Los potenciales de EAM han sido utilizados con éxito en cálculos de fronteras de grano [Vot96]. Hay autores que afirman que son indispensables para poder obtener resultados útiles. Nosotros estamos interesados en estudiar este punto, por lo que buscamos un potencial de tipo EAM simple pero que hubiera dado ya resultados positivos. En esta búsqueda encontramos el potencial de YYYe-Cai [Ye96] (sección 7.6.1). Este potencial está definido en forma analítica, lo que lo hace particularmente fácil de aplicar a simulaciones de dinámica molecular.

Usaremos el potencial de YYYe-Cai para hacer un proceso de enfriamiento igual al de los potenciales de Bristowe y Rydberg para poder comparar los resultados. Esperamos que una simulación con este potencial requiera más tiempo de cálculo, y que tenga resultados cualitativamente equivalentes a Bristowe y Rydberg, pero con valores cuantitativamente más cercanos a los experimentales. Lo que no sabemos es si la mejora en estos valores justifique su uso.

Para el cálculo de las interacciones utilizamos listas de vecinos para cada átomo con radios de corte interno y externo. Cada átomo tiene una lista de sus vecinos más cercanos, a diferencia de métodos en los que la pareja de vecinos sólo se almacena una vez. Nuestro método requiere de más memoria pero tiene mayor flexibilidad en el cálculo de las interacciones. El método lo detallo en la sección 8.8, donde defino los radios de corte interno y externo, quienes determinan la distancia a la que se consideran vecinos y la distancia a la que se consideran las interacciones.

Hicimos algunas pruebas del tamaño adecuado de los radios de corte, finalmente utilizamos para todas las simulaciones los mismos radios de corte que garantizaran una conservación de la energía adecuada mediante actualizaciones de las listas de vecinos frecuentes. Sabemos que estos parámetros no minimizan el tiempo de ejecución pues no aprovechan al máximo las posibilidades de las listas de vecinos, pero si representan un ahorro respecto de los cálculos de todos contra todos. Utilizamos radios de corte internos de 3 unidades reducidas y externos de 4 unidades reducidas. Haciendo actualizaciones con un factor de 5 (ver sección 8.8). Un átomo promedio tiene alrededor de 370 vecinos (dependiendo de la frontera y traslación específica), con alrededor de 165 dentro del radio interno (de interacción).

11.1.6 Condiciones de periodicidad

Decidimos utilizar condiciones de frontera periódicas (sección 11.1.2). La simulación de fronteras de grano parece estar dividida entre utilizar condiciones periódicas y condiciones fijas, aunque recientemente las fronteras fijas parecen estar ganando terreno debido al creciente interés por estudiar fronteras de periodicidad grande [Rom99].

Como explico en la sección 6.3 el uso de condiciones de frontera fijas induce violaciones en las leyes de conservación. En los primeros experimentos buscaremos no sólo observar el comportamiento de las fronteras, sino el del programa, por lo que preferimos limitarnos a las violaciones causadas por el algoritmo de integración y de temperatura constante.

Utilizamos condiciones de frontera periódicas en 3 dimensiones, lo que implica que además de la frontera GCSN el sistema contiene una frontera del tipo que generan los modelos anteriores (que afectuosamente llamamos *la competencia*). Dependiendo del alcance del potencial y del tamaño de la celda puede haber interferencia entre ambos tipos de fronteras, lo que debemos considerar en el análisis de los resultados.

Podemos determinar lo que ocurre con los átomos de una frontera GCSN durante el proceso de enfriamiento y observar sus diferencias en un diagrama bidimensional, sin embargo para las fronteras *de la competencia* no podemos hacer lo mismo porque no están definidos cuales átomos son frontera. Lo normal en estos casos es hacer un diagrama de desplazamientos para todos los átomos de la celda.

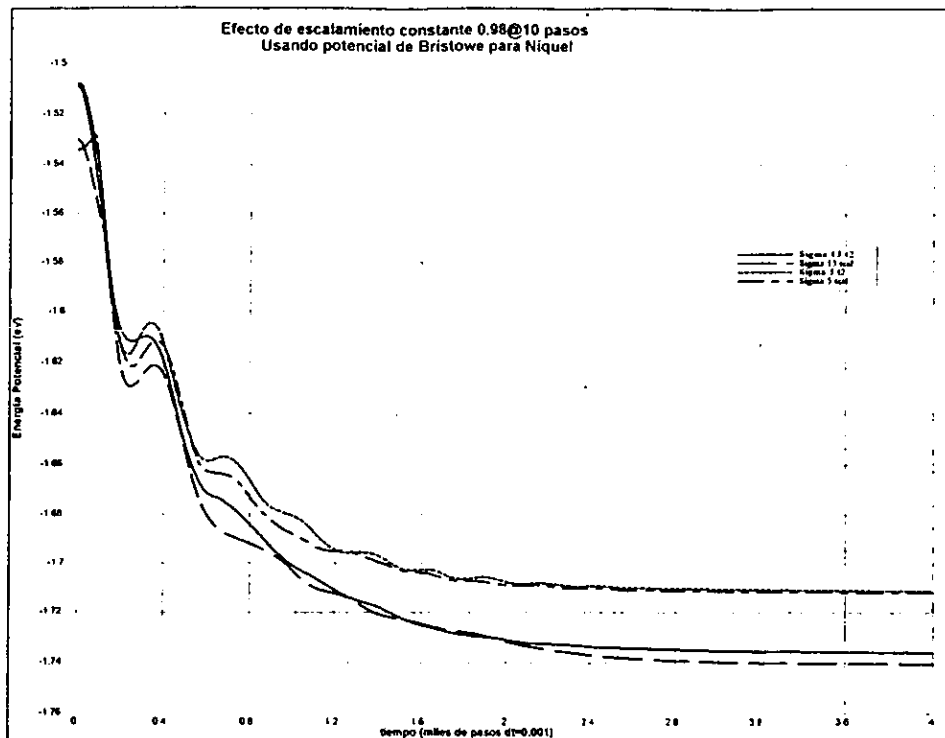


Figura 11-3 Evolución de la energía total durante el enfriamiento de las diferentes fronteras usando el potencial de Bristowe para níquel.

11.1.7 Método de integración

Elegimos el método de integración de Verlet velocidad (sección 6.2), con $dt = 0.001$ por su confiabilidad y simplicidad. En las pruebas de conservación de la energía pudimos observar que esta combinación mantuvo la energía constante con variaciones menores al 1% (y en ocasiones menores), lo que consideramos suficiente dado que para la mayoría de los potenciales no buscamos resultados cuantitativamente exactos.

11.2 Resultados

Una simulación de dinámica molecular, como cualquier experimento, da lugar a una gran cantidad de resultados, y nunca es posible extraer de ella toda la información. Planteados nuestros intereses en la sección 11.1.1 podemos estudiar los resultados en esta dirección. A continuación presento gráficas que permiten dar respuestas a las preguntas.

Las primeras tres figuras 11-3, 11-4 y 11-5 muestran la evolución de la energía total por átomo durante las simulaciones de enfriamiento. Para los potenciales de pares se muestran las cuatro fronteras utilizadas: $\Sigma 13t2$, $\Sigma 13tesl$, $\Sigma 5t2$ y $\Sigma 5tesl$, mientras que para el potencial de tipo EAM sólo se muestra $\Sigma 13t2$. Recordemos que todas las fronteras fueron simuladas bajo los mismos parámetros: evolución clásica NVE durante 10 pasos, luego de los cuales se reducía la temperatura en un 2% (escalamiento por 0.98), repitiéndose el proceso 400 veces.

Las siguientes tres figuras 11-6, 11-7 y 11-8 muestran la evolución de la temperatura (en $^{\circ}K$) para los mismos casos que las energías totales. Las curvas mostradas aquí pueden derivarse para encontrar las variaciones en la temperatura (enfriamientos y calentamientos) que es lo que se muestra en las figuras 11-9, 11-10 y 11-11.

Las figuras 11-12, 11-13, 11-14, 11-15, 11-16, 11-17, 11-18, 11-19 y 11-20 muestran las posiciones atómicas. En cada una de las figuras se muestra arriba y a la izquierda las posiciones de los átomos de la frontera según es generada por el modelo GCSN. Arriba a la derecha se muestran las posiciones de los átomos al término del proceso de minimización, mientras que arriba al centro están ambas superpuestas para permitir una fácil comparación. Todas

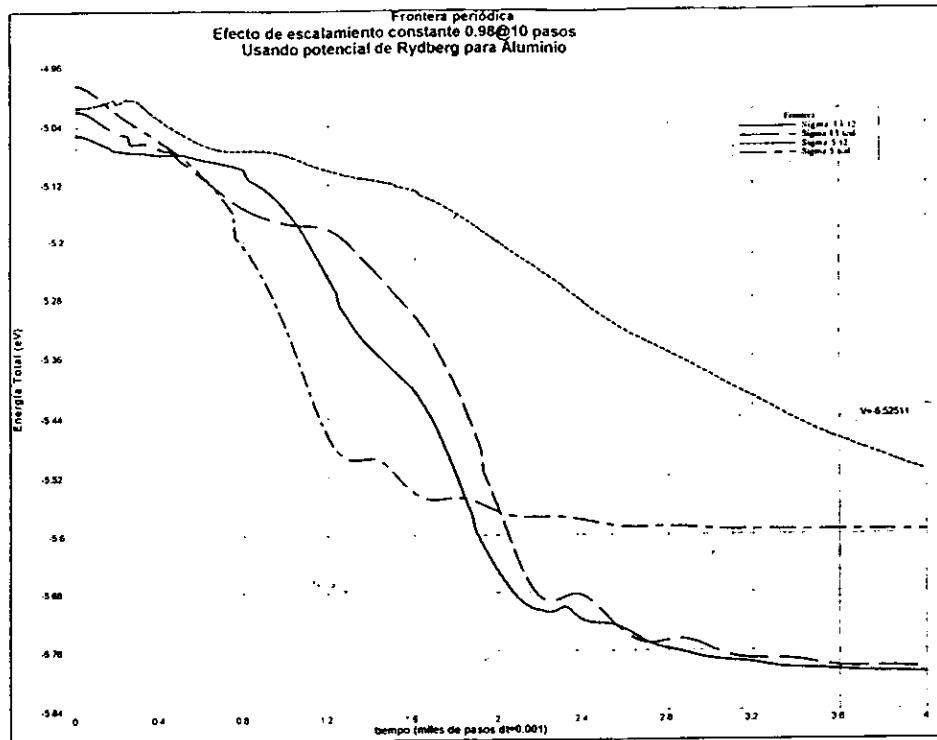


Figura 11-4 Evolución de la energía total durante el enfriamiento de las diferentes fronteras usando el potencial de Rydberg para Aluminio.

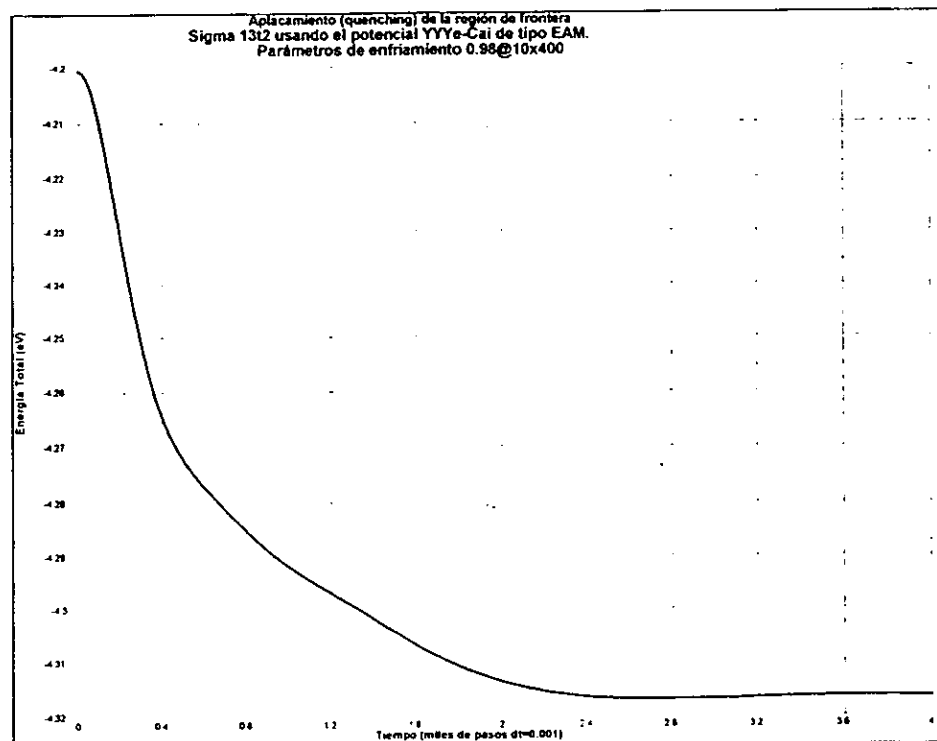


Figura 11-5 Evolución de la energía total durante el enfriamiento de las diferentes fronteras usando el potencial de YYe-Cai de tipo EAM para oro.

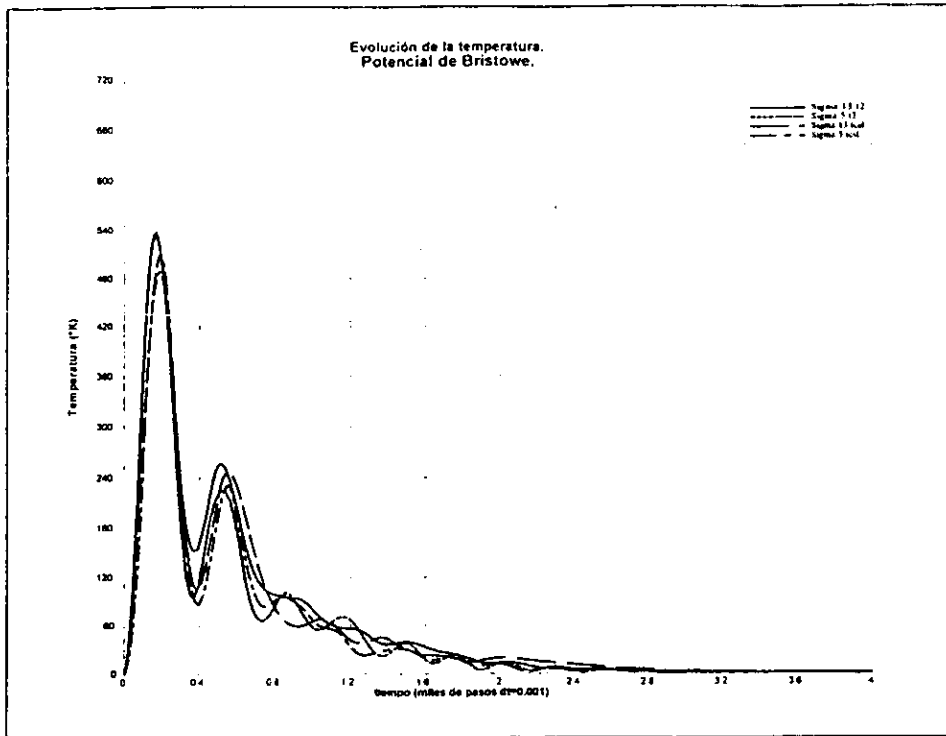


Figura 11-6 Evolución de la temperatura de las diferentes fronteras usando el potencial de Bristowe para Níquel.

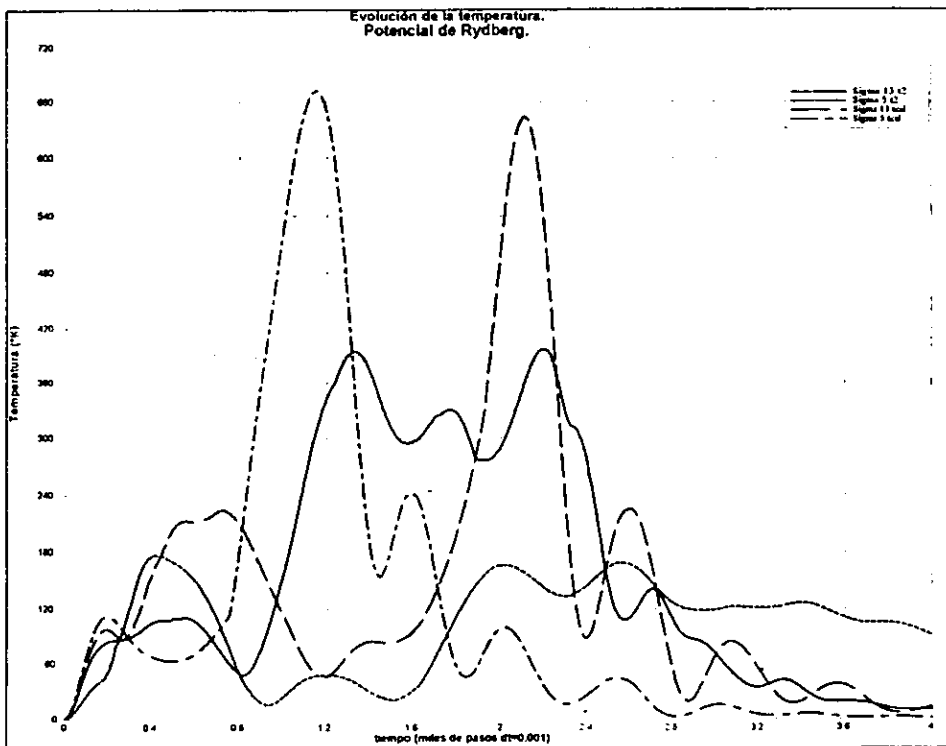


Figura 11-7 Evolución de la temperatura de las diferentes fronteras usando el potencial de Rydberg para Aluminio.

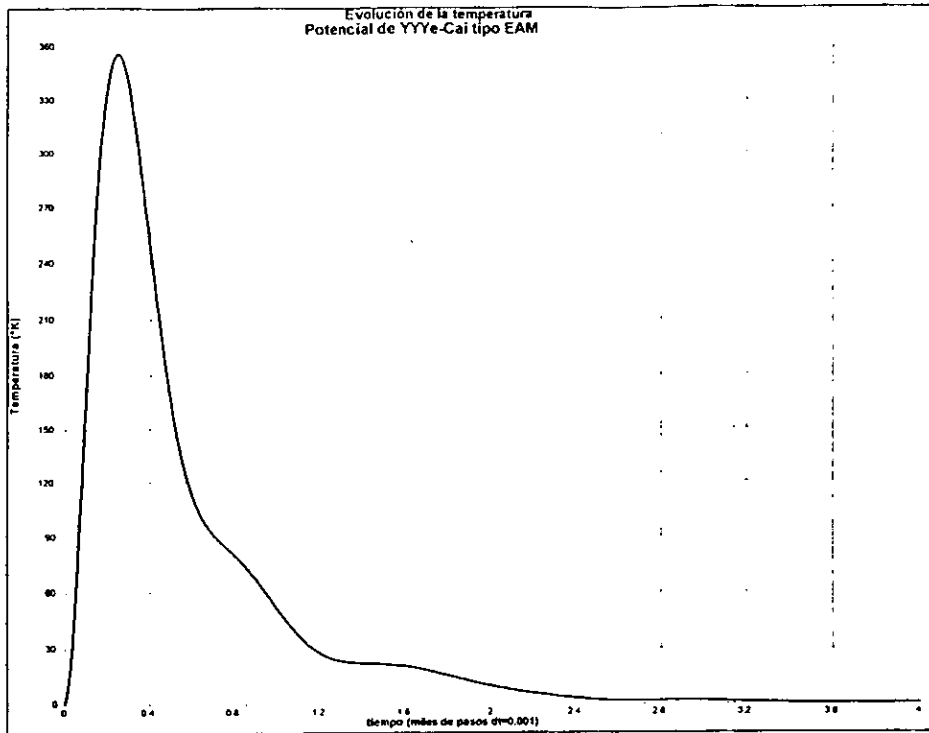


Figura 11-8 Evolución de la temperatura de las diferentes fronteras usando el potencial de YYe-Cai de tipo EAM para oro.

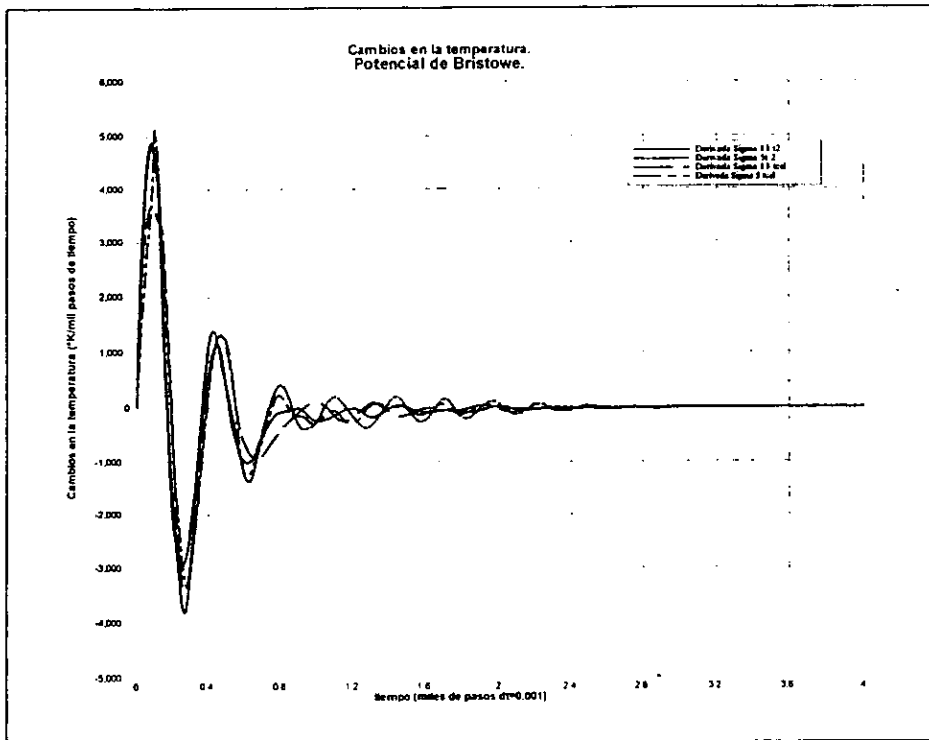


Figura 11-9 Cambios en la temperatura (calentamientos y enfriamientos) de las diferentes fronteras usando el potencial de Bristowe para Níquel.

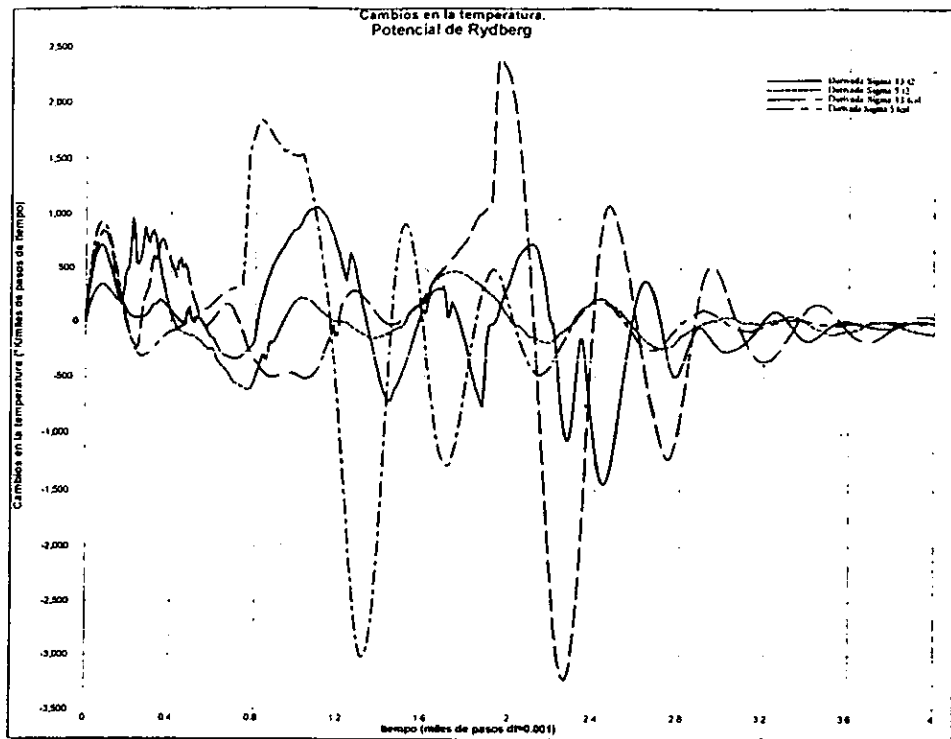


Figura 11-10 Cambios en la temperatura (calentamientos y enfriamientos) de las diferentes fronteras usando el potencial de Rydberg para Aluminio.

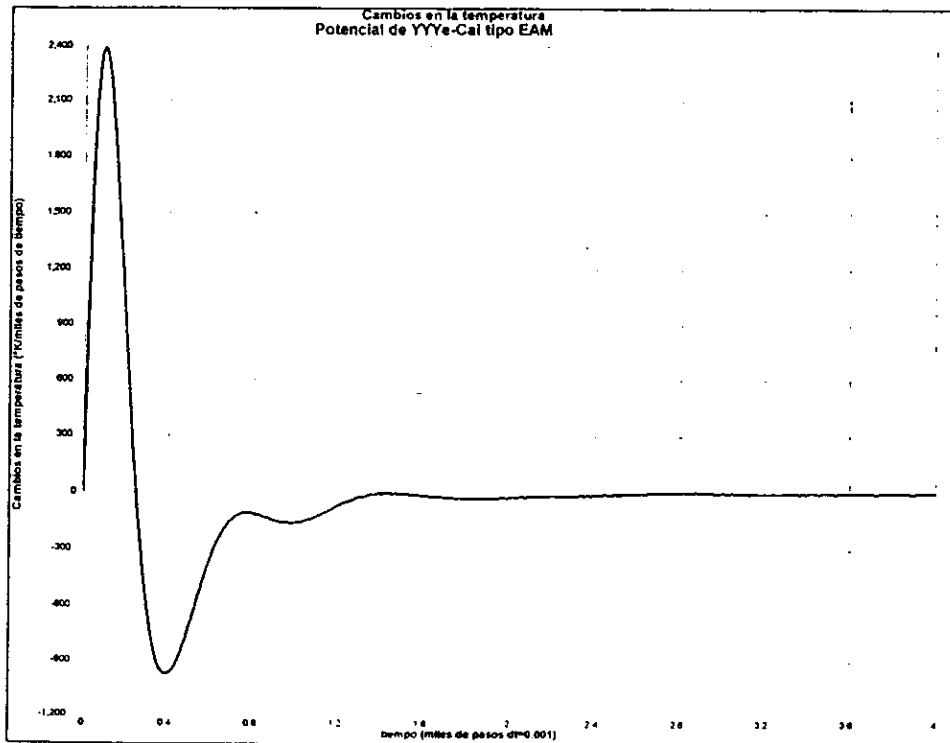


Figura 11-11 Cambios en la temperatura (calentamientos y enfriamientos) de las diferentes fronteras usando el potencial de YYYe-Cal de tipo EAM para oro.

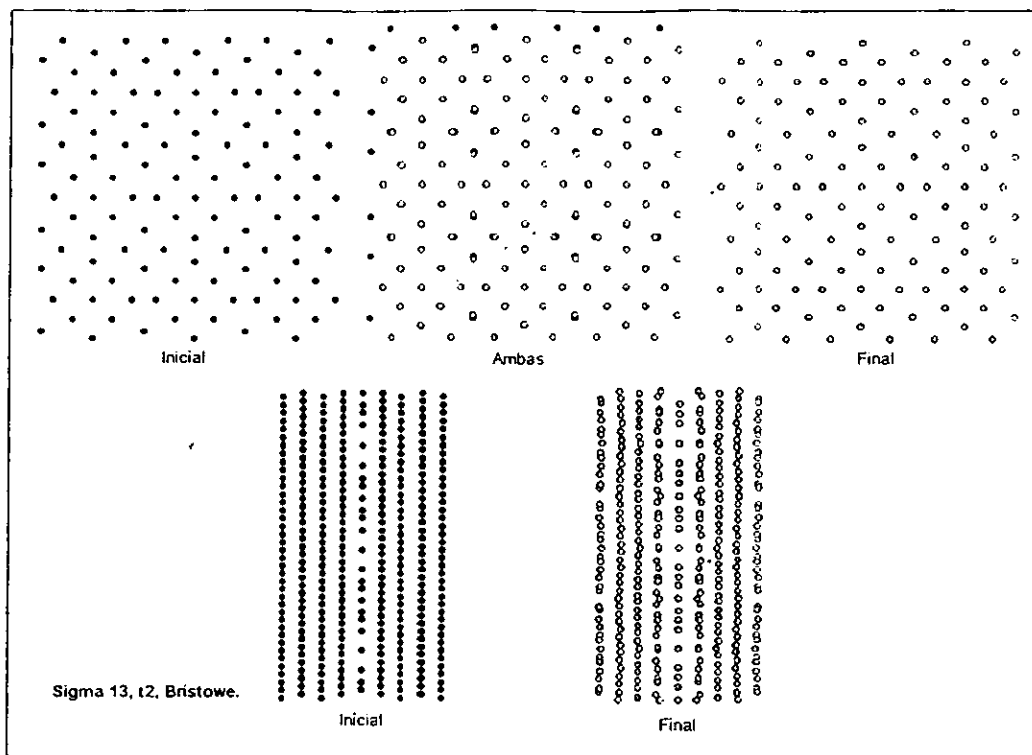


Figura 11-12 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 13t2$ obtenida usando el potencial de Bristowe para Níquel. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

las vistas de los átomos de la frontera son desde una dirección perpendicular al plano de la frontera.

Los diagramas de la parte inferior de las figuras 11-12, 11-13, 11-14, 11-15, 11-16, 11-17, 11-18, 11-19 y 11-20 muestran vistas laterales de los átomos de la celda principal, incluyendo no sólo los átomos de la frontera, sino los de los cristales. La frontera en estos casos es el plano atómico vertical del centro, con cuatro planos atómicos cristalinos a cada lado. Para todas las configuraciones los diagramas son en la misma dirección. El diagrama de la izquierda muestra las posiciones atómicas iniciales según son generadas por el modelo GCSN. El diagrama de la derecha muestra las posiciones finales después del proceso de minimización.

La figura 11-21 muestra las distribuciones de pares inicial y final de la frontera $\Sigma 5t2$ simulada con el potencial de Bristowe para Níquel. Por razones de espacio no muestro las del resto de las fronteras, ya que todas ellas muestran distribuciones equivalentes. En la figura 11-21 están considerados los átomos de la frontera y de los cristales, por lo que no es fácil observar cuales de los pares son efecto de la frontera. La figura 11-22 muestra las distribuciones de la diferencia entre considerar los átomos de la frontera y no considerarlos. Para generarlas retiramos los átomos de la frontera y obtenemos las distribuciones de pares (que no muestro porque corresponden a las de los dos cristales FCC), y las resto de las distribuciones de la figura 11-21. El proceso lo aplico a la configuración inicial (directa del modelo GCSN) y a la final.

En la figura 11-23 se muestra la energía total por átomo de la simulación a temperatura constante (en cada una de sus etapas) de la frontera $\Sigma 13t2$. El potencial utilizado es nuevamente el de Bristowe para Níquel. La figura 11-24 muestra las temperaturas de la misma simulación.

La figura 11-25 muestra las configuraciones inicial y finales de cada temperatura de la simulación de las figuras 11-23 y 11-24. Para cada una de las simulaciones se muestran los átomos en una vista frontal (en una dirección perpendicular al plano de la frontera) y en una vista lateral (sobre el plano de la frontera).

La figura 11-26 muestra vistas frontales obtenidas durante una evolución NVE de la frontera $\Sigma 13t2$ usando el potencial de Lemard-Jones para Aluminio. Las configuraciones fueron tomadas a intervalos de tiempo regulares, mostrandose en la columna de la izquierda la configuración inicial (repetida cinco veces), en la columna de la derecha la configuración instantánea y en la columna del centro ambas para permitir la comparación. La evolución de la energía es equivalente a la parte clásica mostrada en la figura 11-1, aunque de más larga duración.

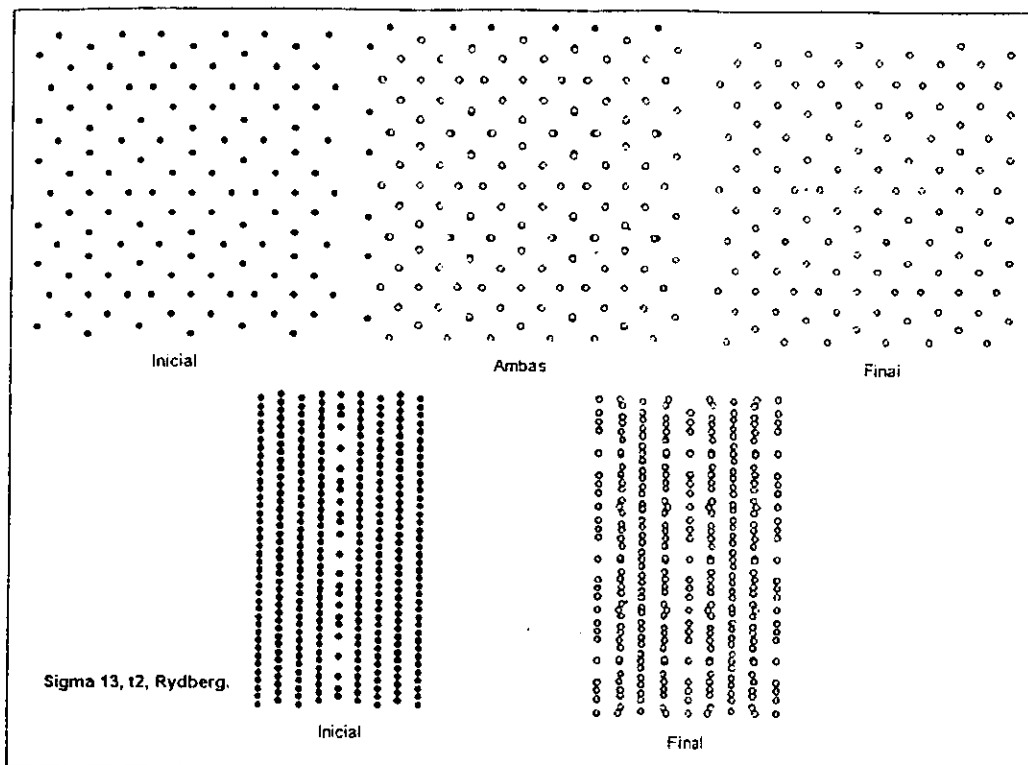


Figura 11-13 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 13t2$ obtenida usando el potencial de Rydberg para Aluminio. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

11.2.1 Análisis de los resultados

De la estabilidad y configuraciones de mínima energía

Las gráficas de las figuras 11-3, 11-4 y 11-5 muestran la energía total por átomo de las diferentes fronteras utilizando los potenciales de Bristowe para Níquel, Rydberg para Aluminio y YYYe-Cai para Oro respectivamente. Salvo en el caso de $\Sigma 5t2$ las fronteras llegaron a una configuración de mínima energía local. Por supuesto no podemos saber que tan bajo es este mínimo con respecto a otros que pudieran existir, pero para esto completamos con los experimentos de temperatura constante. El mecanismo de enfriamiento utilizado permite buscar el mínimo entre más configuraciones que un gradiente conjugado, que otras personas han utilizado con éxito[Cha61].

La primera pregunta que queríamos responder era: ¿son las fronteras generadas por el modelo GCSN estables?. Si observamos las figuras 11-9, 11-10 y 11-11 notamos que cuando se deja al sistema avanzar en el tiempo durante varios pasos a energía constante, la temperatura no aumenta, lo que nos indica que el sistema se encuentra cercano a un mínimo en la energía potencial. De haber elegido un método de minimización por gradiente conjugado no podríamos obtener esta información porque la energía cinética es eliminada a cada paso de integración (sección 6.11.3).

Observamos que en todos los casos $\Sigma 13$ tiene menor energía mínima que $\Sigma 5$, manteniéndose la diferencia inicial en las configuraciones GCSN. Para el caso del potencial de Bristowe la diferencia en energías potenciales iniciales entre las fronteras $\Sigma 13$ y $\Sigma 5$ se mantiene luego de la minimización, pero para el caso del potencial de Rydberg esta diferencia aumenta. Es interesante notar que las fronteras $\Sigma 13tcs1$ y $\Sigma 5tcs1$ tienen energías muy parecidas en una configuración GCSN, sin embargo luego del proceso de minimización la diferencia es porcentualmente mucho mayor.

Según Bristowe (sección 11.1.2) las fronteras de tipo csl deben ser de menor energía que las de tipo 2, aunque no siempre en forma importante. En los experimentos que nosotros hicimos encontramos exactamente eso. En todos los casos, salvo $\Sigma 13$ con Rydberg encontramos que csl es de menor energía que 2. Para $\Sigma 13$ con Rydberg encontramos lo opuesto, pero por muy pequeño margen.

Con las figuras 11-12, 11-13, 11-14, 11-15, 11-16, 11-17, 11-18, 11-19 y 11-20 podemos comparar las configuraciones de mínima energía encontradas con las iniciales generadas por el modelo GCSN. Las posiciones son muy parecidas, independientemente de la frontera o potencial que utilizamos. Notamos que existen diferencias en las posiciones

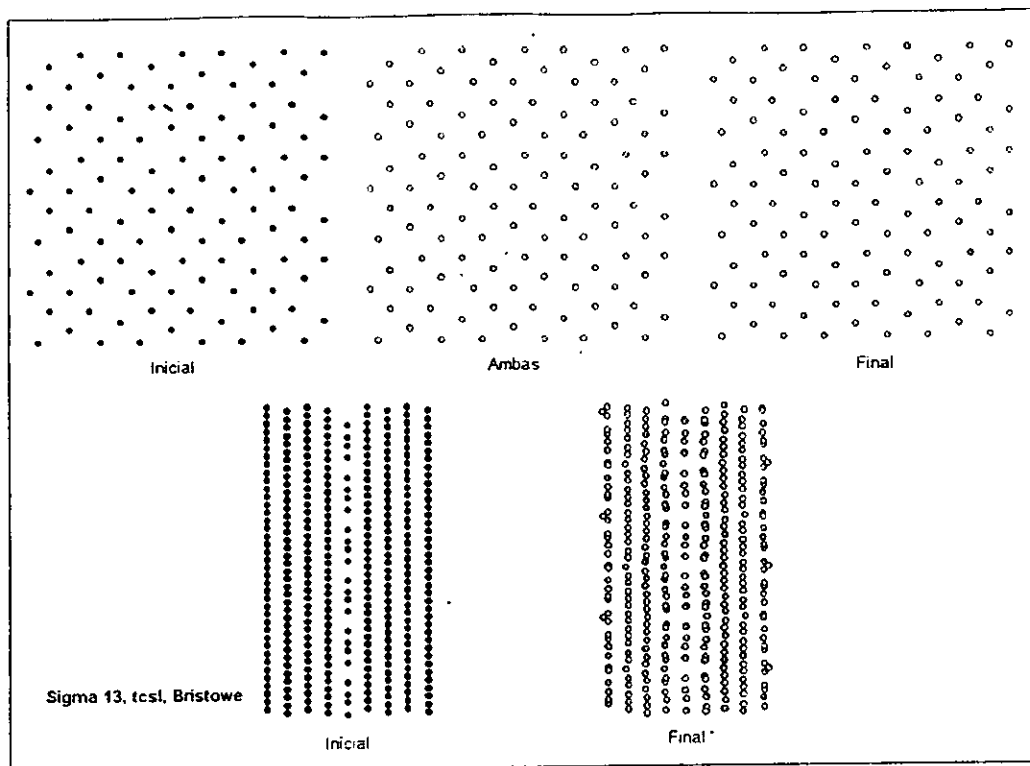


Figura 11-14 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 13tcs1$ obtenida usando el potencial de Bristowe para Níquel. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

atómicas, sin embargo estas son pequeñas respecto a las distancias interatómicas, según se puede observar en la figura 11-21 para una frontera representativa.

En todos los casos observamos que los motivos presentes en las fronteras GCSN se mantienen durante el proceso de minimización y se observan claramente en las configuraciones finales.

En las comparaciones de estructura atómica se puede notar que los átomos que tienen máximo *encaje* (posiciones de coincidencia entre los cristales) se han mantenido en su posición inicial o se han desplazado en forma imperceptible, mientras que los átomos de mínimo *encaje* se han desplazado mucho más. Esto implica que las regiones de baja y alta energía en la frontera están localizadas según lo predicho por el modelo GCSN.

Cuando han ocurrido desplazamientos notables de algunos átomos podemos observar que ocurren manteniendo el motivo y la simetría de la frontera. En fronteras como la $\Sigma 5t2$ con el potencial de Rydberg (figura 11-17) encontramos mayores diferencias con la configuración inicial, sin embargo podemos explicar esto por de diferentes maneras: 1) Aún no se ha llegado a la configuración de mínima energía, lo que podemos observar en la gráfica de enfriamiento (figura 11-4). Para poder comparar con el resto de las fronteras decidimos utilizar los mismos parámetros de enfriamiento, sin embargo observamos un comportamiento notablemente distinto de esta frontera, y que valdría la pena estudiar más adelante. 2) Si realizamos un desplazamiento de los átomos de la configuración inicial podemos encontrar un mejor encaje, esto puede significar que ocurrió un desplazamiento del sistema, lo cual no está prohibido por el tipo de simulación (siempre y cuando el momento total se mantenga constante igual a cero).

En las vistas laterales podemos observar los desplazamientos en direcciones perpendiculares a la frontera. Observamos que estos son muy pequeños, y cuando ocurren se mantiene la simetría del sistema. La vista lateral de la configuración inicial muestra claramente los diferentes planos de cada cristal (cuatro) y el plano de la frontera. En las configuraciones finales pueden distinguirse los planos, pero no en todas con la misma claridad, por ejemplo $\Sigma 5t2$ con Bristowe mantiene los planos notablemente mejor que $\Sigma 5t2$ con Rydberg. Si en las figuras pudieran mostrarse con claridad las diferentes especies atómicas se podrían distinguir mejor los planos, ya que aún cuando se han movido los átomos no ocurren intercambios de posiciones entre especies.

Que los motivos y las simetrías se mantengan habla no sólo de las fronteras GCSN, sino del programa y de las condiciones de frontera utilizadas. Significa que es probable que el programa sea correcto en este sentido.

No debemos olvidar que las condiciones de frontera son periódicas, lo que implica que los planos más lejanos de

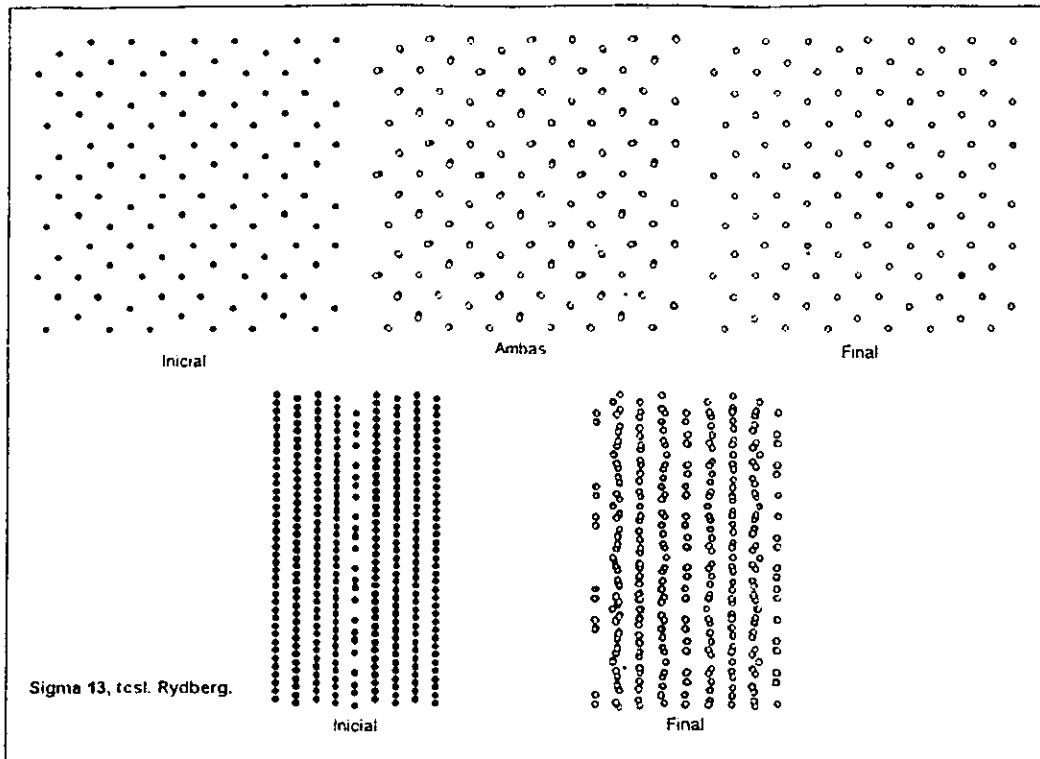


Figura 11-15 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 13tcs1$ obtenida usando el potencial de Rydberg para Aluminio. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

la frontera de cada cristal se encuentran en interacción entre si, esto implica que al mismo tiempo de simular una frontera GCSN estamos simulando una frontera tradicional (o *de la competencia* como dijimos antes). Debido al alcance del potencial y al número de planos cristalinos utilizados sabemos que existe una interacción entre ambos tipos de fronteras, esto puede influir en la configuración de equilibrio y "juega" en contra de nuestras intenciones. Sería conveniente repetir los experimentos con más planos atómicos, de tal manera que pudieramos reducir a un mínimo la interacción entre las fronteras.

A pesar de la interacción entre las fronteras podemos notar que siempre que se ha mantenido el orden en la configuración este parece ser mayor en la región de la frontera GCSN. Los planos de la región de frontera *competencia* se pierden en muchos casos, mientras que los planos de la frontera GCSN se mantienen, incluso en las fronteras de más cambio.

Para el caso del enfriamiento utilizando el potencial de YYe-Cai de tipo EAM no encontramos diferencias perceptibles en la configuración final con respecto a la inicial o a las finales encontradas con otros potenciales. Este es uno de los resultados principales que queríamos encontrar. Para una frontera $\Sigma 13t2$ enfriada con este proceso sabemos que no es necesario pagar el costo computacional de utilizar un potencial de tipo EAM. Es necesario repetir el experimento con otras fronteras y otros procesos de enfriamiento, sin embargo la existencia de un caso en el cual no parece haber ninguna ventaja en usar el potencial EAM es importante. También es importante porque permite observar el funcionamiento del programa al utilizar este tipo de potenciales.

De la temperatura de fusión

Mediante la simulación de enfriamiento de las figuras 11-3, 11-6 y 11-9, encontramos la configuración a $0^\circ K$ de la frontera $\Sigma 13t2$ con el potencial de Bristowe para Níquel. Para verificar la estabilidad a temperaturas mayores utilizamos una simulación de temperatura constante, de la cual mostramos las figuras 11-23 y 11-24. En esta simulación el sistema se mantenía a una temperatura constante durante un tiempo mayor al de equilibrio. Al finalizar este tiempo se obtenía su configuración y se llevaba a una nueva temperatura para repetir el proceso. En las figuras podemos observar el comportamiento de la energía total y la temperatura del sistema completo.

Observando las fluctuaciones en la temperatura podemos aproximar la duración del período de equilibrio, que en

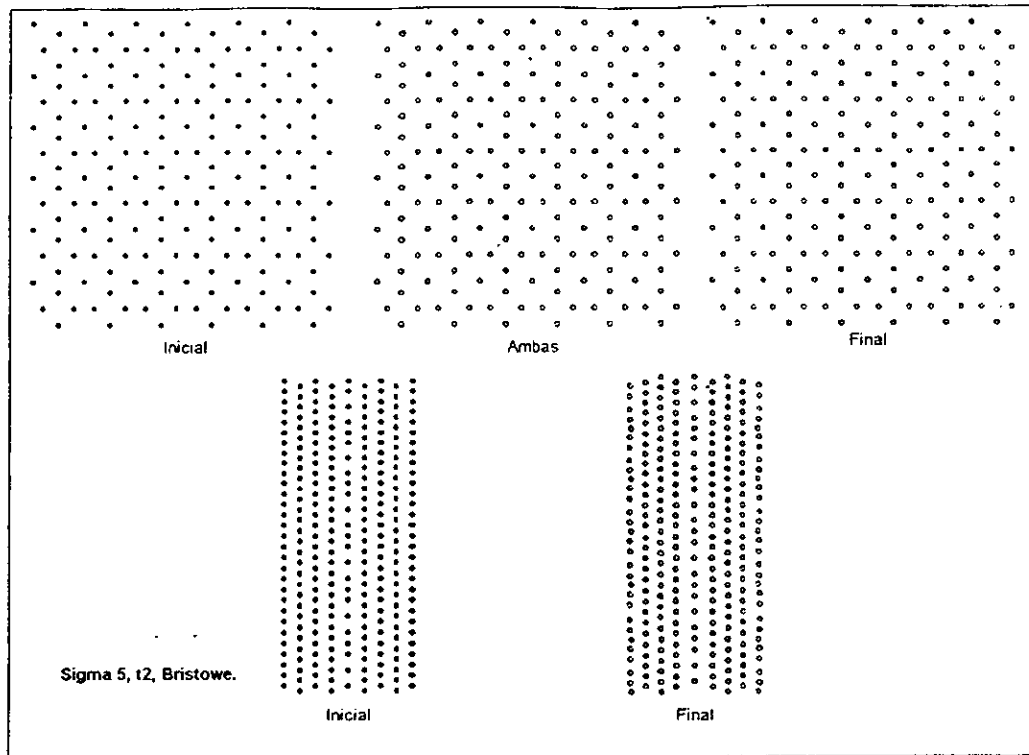


Figura 11-16 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 5t2$ obtenida usando el potencial de Bristowe para Niquel. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

todos los casos resulta varias veces menor a la duración de la etapa. De esta forma suponemos que las mediciones hechas al final de cada etapa corresponden a un estado de equilibrio. Las fluctuaciones en la energía total asociadas con mantener constante la temperatura y las fluctuaciones aceptadas en la temperatura son mayores conforme aumenta la temperatura, lo cual coincide con lo esperado y nos induce a creer que el proceso realizado por el programa es correcto.

En la figura 11-26 observamos las configuraciones al final de cada etapa de temperatura constante. En ellas podemos observar aproximadamente la temperatura a la que se pierde la identidad de los motivos.

Es evidente que a $2000^{\circ}K$ la frontera se ha destruido y los planos atómicos se han difundido. Podemos considerar que en ese estado el material se ha fundido. A temperaturas de $0^{\circ}K$ y $400^{\circ}K$ observamos lo contrario. La configuración se mantiene y la frontera no pierde su identidad. Esto significa que la temperatura de recristalización debe encontrarse entre los 400 y los $2000^{\circ}K$, lo que tiene sentido, pues sabemos que debe ser menor que la temperatura de fusión, que en este caso es de $1726^{\circ}K$.

Observando las configuraciones de frente y los planos atómicos de lado parece ser que la temperatura de recristalización se encuentra entre los 1200 y los $1600^{\circ}K$. Esto sería algo mayor que lo que esperábamos (entre los 800 y los $1200^{\circ}K$), sin embargo no hay que olvidar que la duración de las etapas es relativamente pequeña. Si repitiéramos el experimento simulando mucho más tiempo podríamos encontrar con más precisión la temperatura de recristalización. En la sección 6.7 describo varios métodos para medir la temperatura de fusión de un material, algunos de los cuales pueden ser aplicados en este caso en forma directa.

A $800^{\circ}K$ notamos también que la frontera GCSN parece más estable que la competencia, que ya ha empezado a perder estructura. A $1200^{\circ}K$ definitivamente la frontera de la competencia se ha perdido y el plano de la GCSN es ya difuso.

Durante diferentes simulaciones de prueba habíamos observado configuraciones que nos llamaron la atención, sin embargo estas no las observamos siempre. Un ejemplo de configuración que nos llama la atención es la última mostrada en la figura 11-26. La figura 11-27 es una copia en mayor tamaño. En esta figura se ha marcado con la línea diagonal más gruesa el plano de la simetría de espejo, pero también algunos planos sobre los que los átomos parecen alinearse. En el acomodo de los átomos parecen distinguirse dislocaciones, sin embargo estas pierden su identidad por la proximidad y por el nivel de detalle atómico en que nos encontramos. Harían falta simulaciones

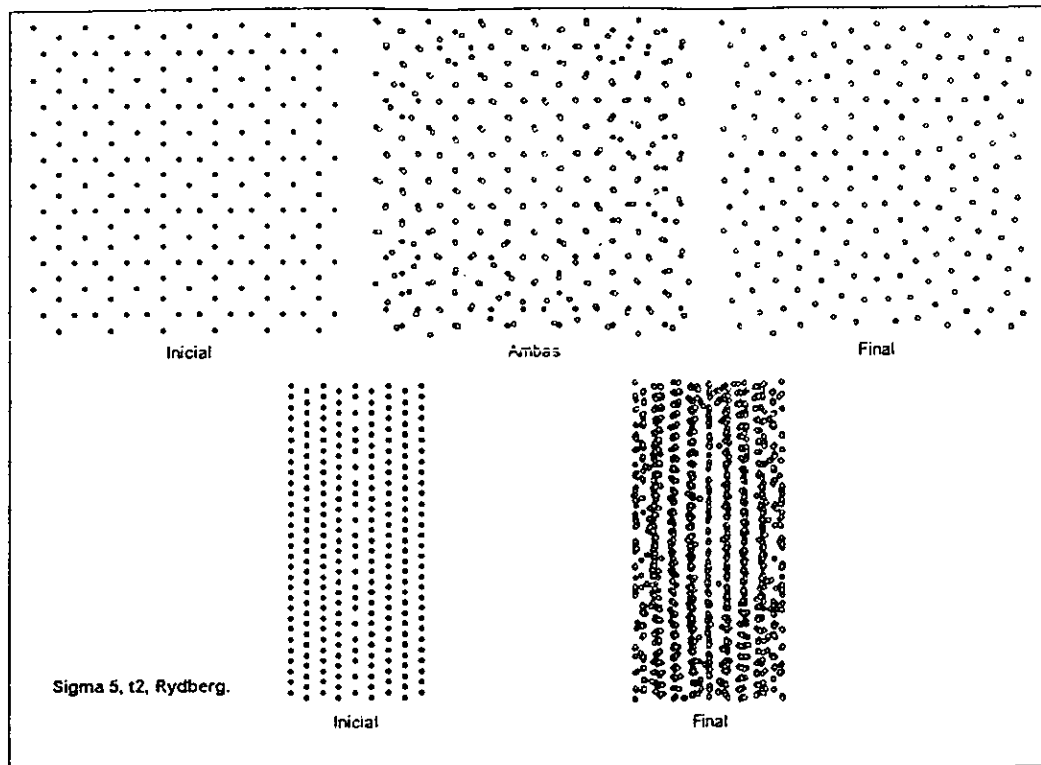


Figura 11-17 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 5/2$ obtenida usando el potencial de Rydberg para Aluminio. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

con muchos más átomos buscando observar este efecto. La simetría la encontramos no sólo en la dirección mostrada sino en la perpendicular, lo que indica que es poco probable que sea un error de la simulación. La razón por la cual esta simetría la observamos sólo en la última configuración de la evolución con energía constante es que la duración de esta la elegimos tal que llegara a lo que buscamos. Hubieramos podido continuar y la configuración se hubiera perdido.

Observamos este tipo de configuraciones en diferentes fronteras y circunstancias. Pensando que existía la posibilidad de que fuera un modo normal de vibración de la frontera tomamos la frontera que se muestra en la figura 11-27 como configuración inicial y la dejamos evolucionar con energía constante y con temperatura constante. En ninguno de los casos pudimos observar que fuera un modo normal de vibración.

11.2.2 Conclusiones de fronteras de grano

Las simulaciones de enfriamiento llevaron configuraciones de mínima energía local muy parecidas a las generadas por el modelo GCSN, lo que permite concluir que bajo los diferentes potenciales probados las fronteras GCSN son de equilibrio y baja energía. Los cálculos con un potencial de tipo EAM dieron configuraciones equivalentes a las de los potenciales de pares.

La energía de las configuraciones de mínima energía encontradas muestra que $\Sigma 13$ tiene menor energía que $\Sigma 5$, lo que coincide con los experimentos y con los cálculos hechos con simulaciones de Monte Carlo [Rom99]. Las fronteras en estado traslacional tcs1 llegaron a configuraciones de menor energía que las en estado t2 para todos menos un caso. Las diferencias de energía entre ambos estados traslacionales son muy pequeñas.

La periodicidad en las configuraciones intermedias y finales indican que el programa maneja en forma adecuada las condiciones de frontera periódicas. Las energías de las configuraciones finales muestran que también se están calculando correctamente las interacciones, comparadas con cálculos teóricos simples.

Usando simulaciones de temperatura constante encontramos que la temperatura de recristalización es de entre

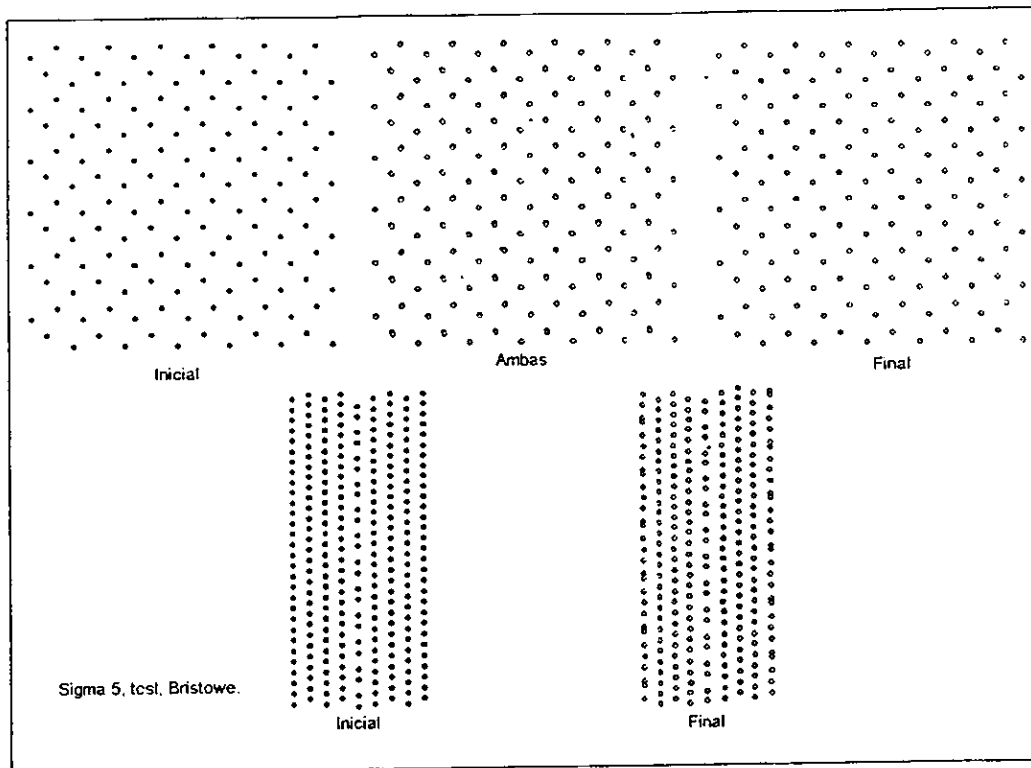


Figura 11-18 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 5_{tcsi}$ obtenida usando el potencial de Bristowe para Níquel. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

1200 y 1600°K, menor a los 1726°K de la temperatura de fusión de un cristal de Níquel. Estas simulaciones permiten observar también que las fronteras GCSN son estables a temperaturas ambiente.

Las curvas de enfriamiento de las diferentes fronteras muestran que los enfriamientos para potenciales *con joroba* son diferentes a los de potenciales *sin joroba*, pero son parecidos al del potencial tipo EAM. Aparentemente para obtener conclusiones cualitativamente correctas no es necesario pagar el precio computacional de usar potenciales de tipo EAM.

En todos los casos encontramos que la frontera GCSN mantiene su estructura a mayores temperaturas que las fronteras *de la competencia*. Las fronteras GCSN mostraron mayor estabilidad, aún cuando encontramos que para el número de átomos utilizado hay interacciones entre ambos tipos de frontera. En ningún caso se destruyó una frontera GCSN mientras se mantuvo una frontera de la competencia.

Los motivos de las fronteras $\Sigma 5$ y $\Sigma 13$ se mantienen después del proceso de minimización de energía, pero también a temperaturas ambiente y hasta llegar al intervalo de la temperatura de recristalización. Las coincidencias entre las configuraciones iniciales y las de mínima energía son notables, encontrando además que los sitios de mayor diferencia son aquellos donde se esperaba un peor *encaje*, mientras que los desplazamientos mínimos son en los puntos de máxima coincidencia que ya conocíamos. Esto implica que las simetrías y las redes de dislocaciones de las fronteras GCSN coinciden con las de las fronteras de mínima energía.

Encontramos estados con simetría de espejo sobre un plano que corta en diagonal (1,1,0) que pueden representar modos normales de vibración. Las películas de simulaciones de energía constante y de temperatura constante parecen indicar lo contrario.

Ningún resultado es opuesto a lo que muestran los experimentos o a lo que esperábamos.

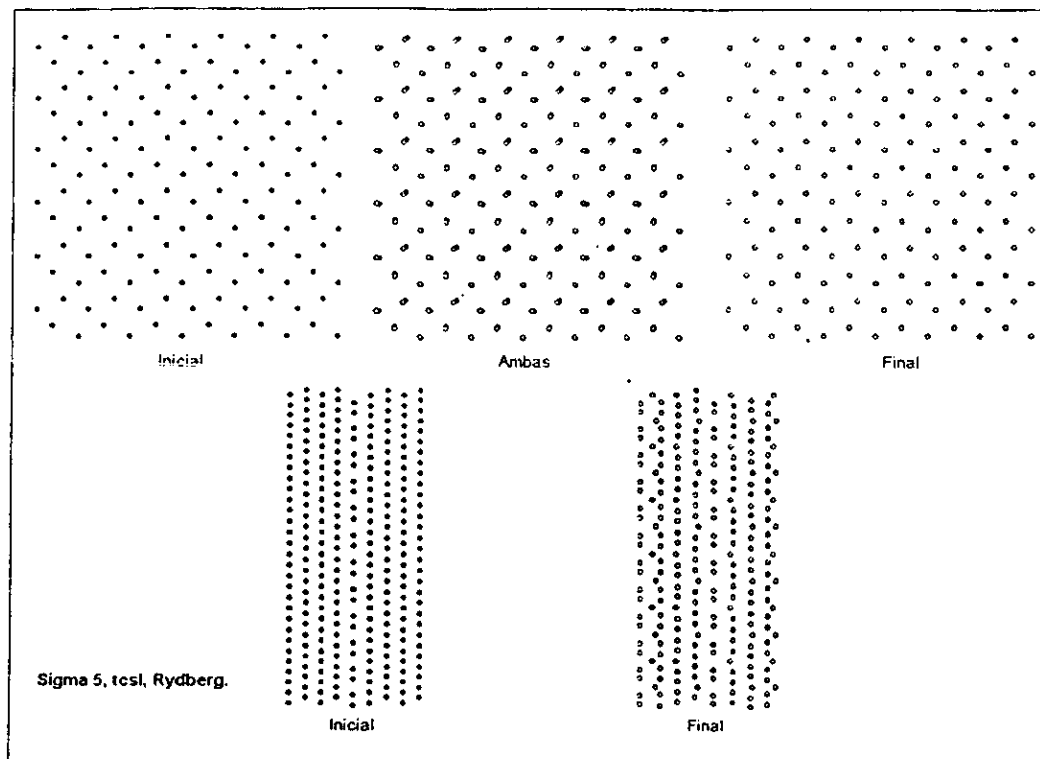


Figura 11-19 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 5_{tcsl}$ obtenida usando el potencial de Rydberg para Aluminio. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

11.3 Conclusiones

En los objetivos de la tesis planteamos tres objetivos: estudiar fronteras de grano GCSN, generar un programa de dinámica molecular y generar un manual de fronteras de grano y dinámica molecular. Para concluir y evaluar los logros los revisare en respecto a los tres objetivos.

- **Del programa** - Los resultados de las fronteras muestran un programa que reproduce en forma adecuada los diferentes aspectos esenciales de la dinámica molecular. Las condiciones de frontera son correctas, las interacciones y el algoritmo de integración, la lectura y escritura de datos, el manejo de la memoria y los demás aspectos discutidos en la sección 9 funcionan como esperamos.

El hecho de que el programa se encuentre en objetos permite su fácil revisión, modificación y extensión. Es plausible adaptarlo en el futuro para simulaciones mixtas con Monte Carlo, así como agregarle las funciones que se consideren necesarias. Por haber utilizado ANSI C++ en el 99.9% de las líneas es fácilmente portable a otras máquinas del IFUNAM.

La forma en que se le dan instrucciones por medio del archivo ini es algo complicada, pero simple en comparación con la mayoría de los programas de dinámica molecular.

Para los cálculos de fronteras de grano hechos para esta tesis el programa resulto más que adecuado. Corresponde a los futuros usuarios determinar si la manera en que se diseño y construyo son en realidad correctas.

El programa permitió observar propiedades desconocidas de las fronteras, como el surgimiento de nuevas simetrías (figura 11-27), lo que muestra su utilidad inclusive desde los primeros cálculos.

- **De las fronteras de grano** - Las fronteras GCSN resultaron estables ante todos los potenciales utilizados, independientemente de su forma. Las energías mínimas de cada una de las fronteras y traslaciones corresponde con lo esperado según la teoría y los cálculos previos.

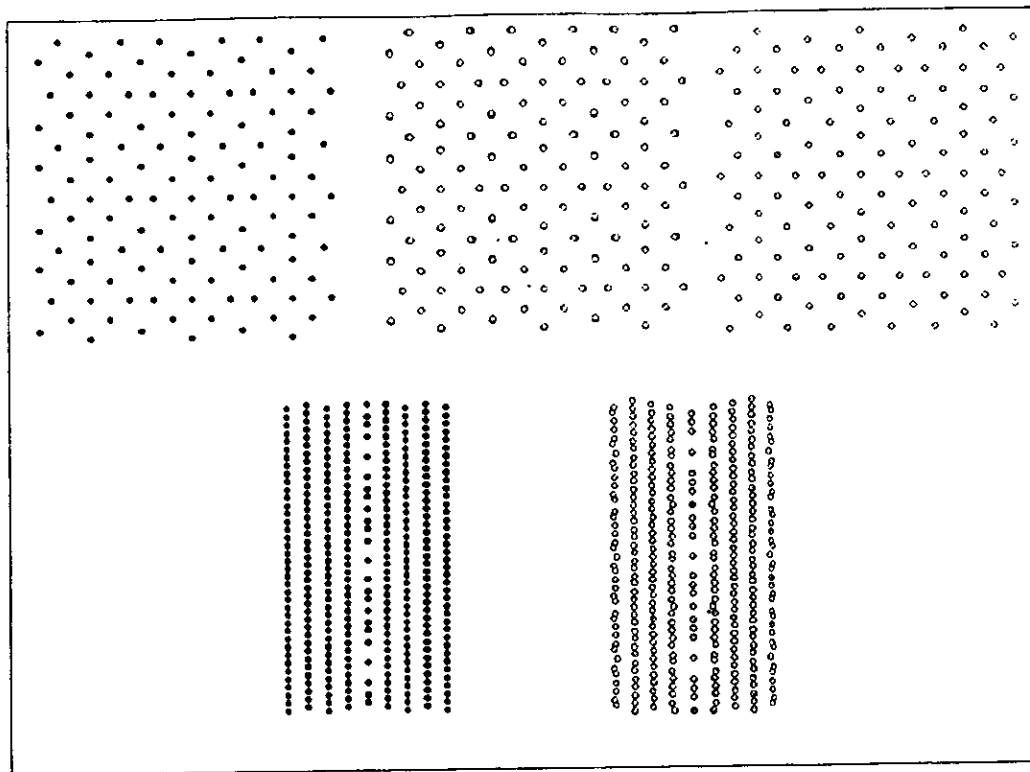


Figura 11-20 Configuraciones inicial, de energía mínima y comparada de la frontera $\Sigma 13t2$ obtenida usando el potencial de YYe-Cai de tipo EAM para oro. Arriba vista frontal de los átomos de la frontera. Abajo vista lateral de los átomos de la celda principal.

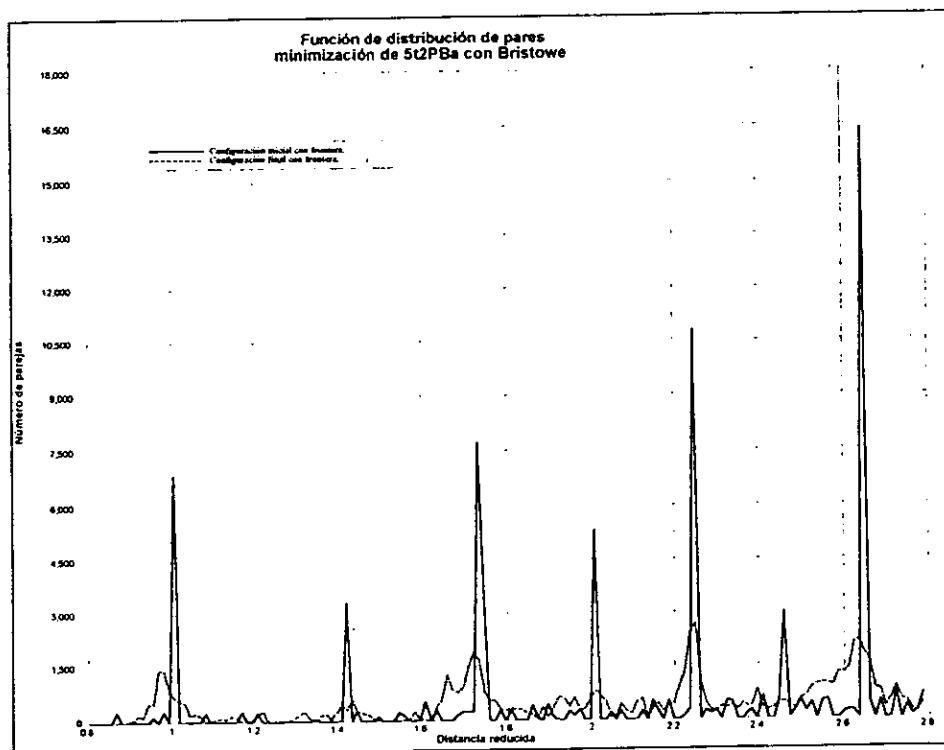


Figura 11-21 Distribuciones de pares entre los átomos de la celda principal en la configuración inicial ($\Sigma 5t2$) y la de mínima energía según un potencial de Bristowe para Níquel.

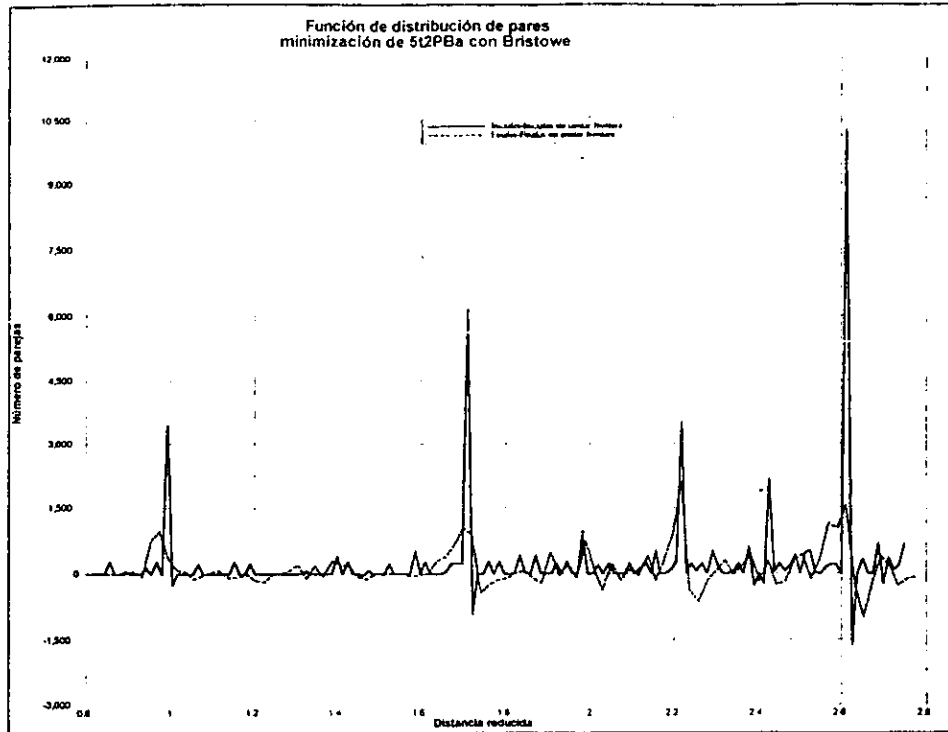


Figura 11-22 Diferencias entre las distribuciones de pares considerando y sin considerar los átomos de la frontera $\Sigma 512$. La figura 11-21 muestra las distribuciones de pares inicial y final.

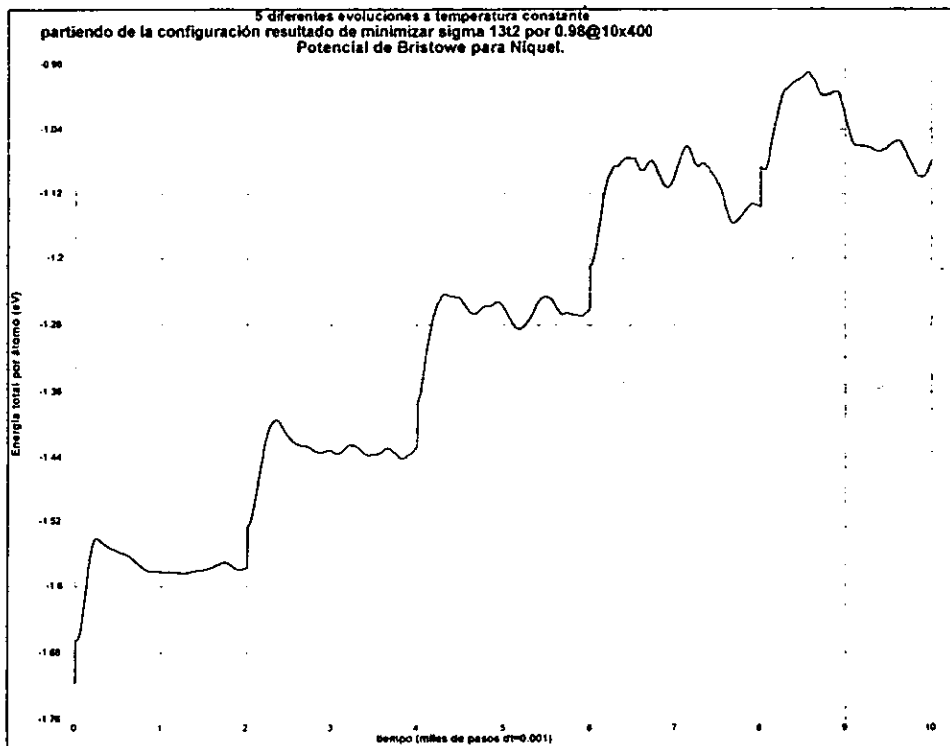


Figura 11-23 Evolución de la energía total durante simulaciones a diferentes temperaturas constantes de $\Sigma 1312$. Notese como las fluctuaciones son mayores para mayores temperaturas. La figura 11-24 muestra la evolución de las temperaturas.

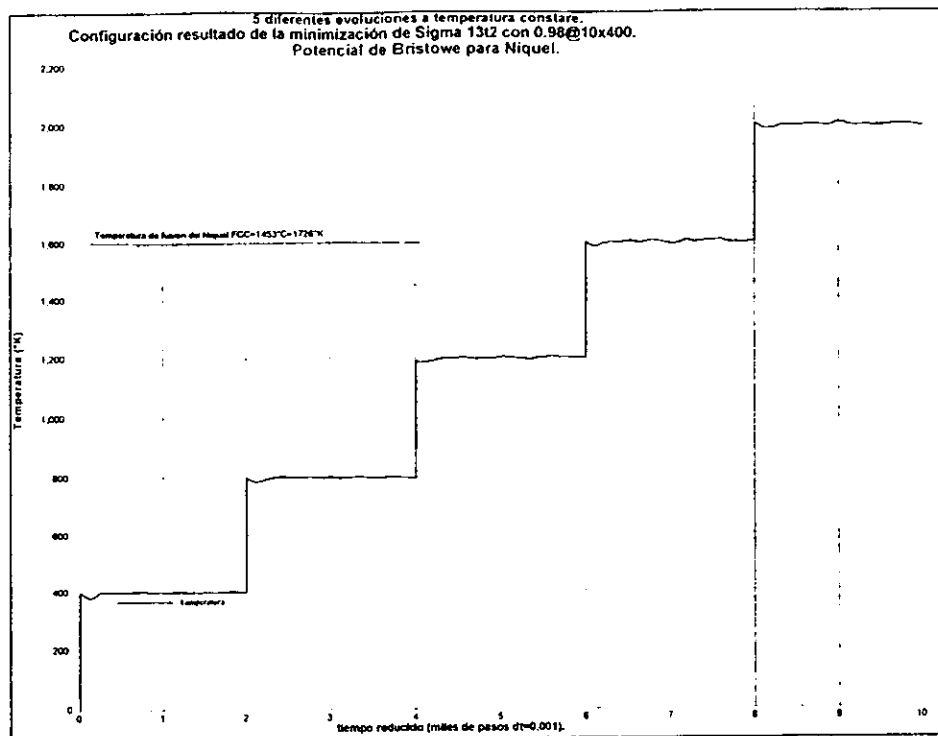


Figura 11-24 Evolución de la temperatura durante simulaciones a diferentes temperaturas constantes. Notese el período de equilibrio al inicio de cada temperatura, así como que las oscilaciones son pequeñas pero ligeramente mayores para mayores temperaturas. La energía total por átomo se muestra en la figura 11-23.

Las simulaciones a temperatura constante mostraron estabilidad a temperaturas mayores que 0°K , delimitando la temperatura de recristalización en el rango esperado.

Todos los resultados obtenidos concuerdan con lo esperado y están en perfecta concordancia con los experimentos, reforzando la confianza en el modelo GCSN para generar fronteras de grano.

Se encontraron más preguntas que respuestas acerca de las fronteras observadas, lo que muestra que el tema se encuentra muy lejos de agotarse, dejando ampliamente justificado el tiempo invertido en estudiar el problema.

Para los cálculos que nosotros hicimos no encontramos justificación de utilizar potenciales de tipo EAM, lo que significa que podemos mantenernos usando potenciales de pares adaptados a fronteras de grano, que aunados al uso del modelo GCSN reducen enormemente el tiempo requerido para estudiar fronteras de grano.

- **Del manual-** El manual de dinámica molecular incluido en este texto contiene los elementos clave para iniciar a cualquier estudiante de los semestres finales de la carrera de física al estudio de la simulación por dinámica molecular. Sin embargo no puede sustituir en ningún momento la experimentación. *Nadie puede aprender a caminar leyendo la teoría.* Serán necesarias muchas horas de práctica y probablemente semanas de programación desperdiciadas antes de poder utilizar lo discutido en estos textos. Espero que la cantidad de tiempo y energía perdidos sean menores gracias a mi trabajo.

En este texto discutimos algunos puntos relacionados con la elección del lenguaje de programación, así como de la programación orientada a objetos, utilizando el programa como ejemplo. Este punto puede resultar sumamente útil para futuros estudiantes, disipando algunas de las dudas de estos respecto a las diferencias entre los distintos lenguajes.

El éxito o fracaso del propósito educativo de este texto no podemos medirlo ahora, es necesario observar la forma en que lo aprovechan los usuarios futuros.

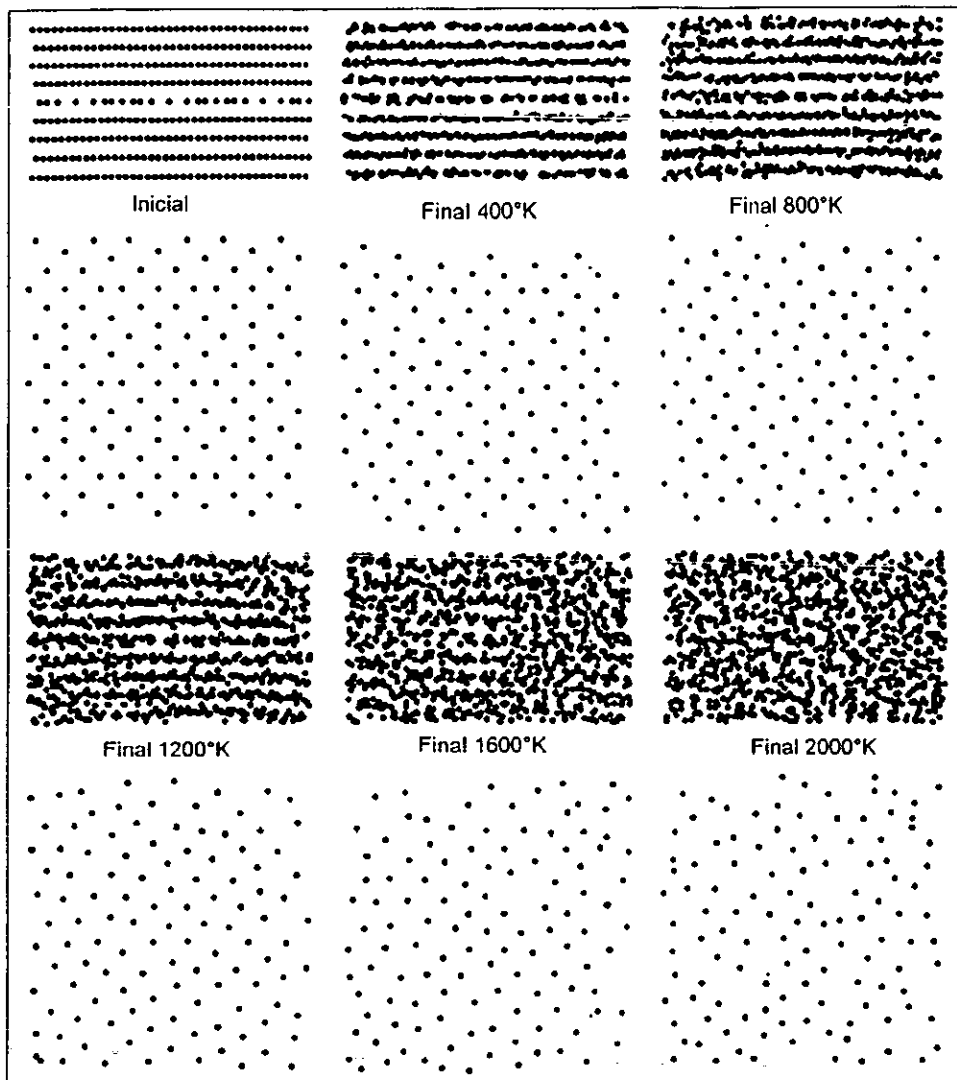


Figura 11-25 Configuraciones frontal (sólo los átomos de la frontera) y lateral (todos los átomos de la celda principal) de las configuraciones finales obtenidas tras simulaciones a temperatura constante usando potencial de Bristowe para Níquel. La temperatura de fusión del Níquel es $1726^{\circ}K$. A simple vista se distingue que por debajo de los $800^{\circ}K$ se mantiene la estructura, y esta se ha perdido por arriba de los $1600^{\circ}K$. Para los $1200^{\circ}K$ haría falta un estudio más cuidadoso. Las figuras 11-23 y 11-24 muestran la energía y temperatura de la simulación respectivamente.

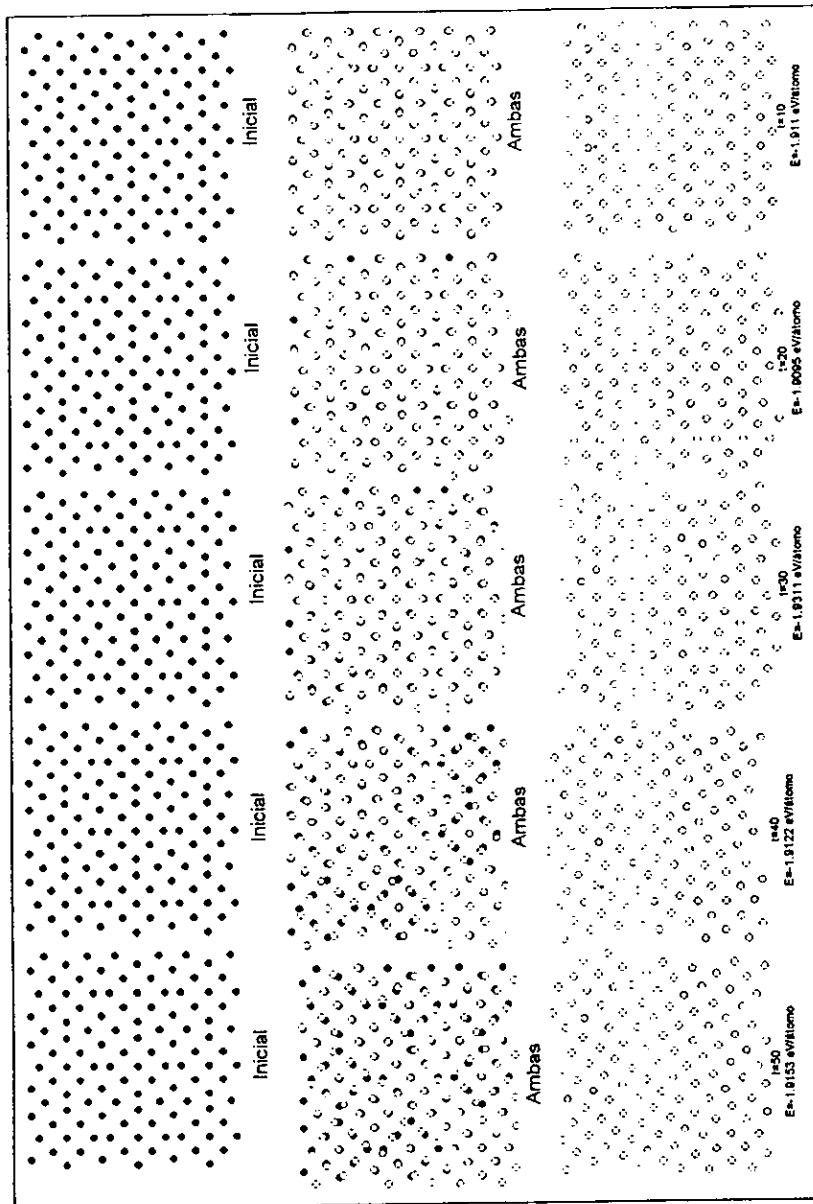


Figura 11-26 Configuraciones intermedias durante una simulación clásica de energía constante. Notes como sin cambios notables en las energías se observa una simetría interesante en la última configuración. La frontera es $\Sigma 13\bar{2}$ y el potencial es Lennard-Jones para aluminio. En la figura 11-1 se muestra un ejemplo de la evolución de la energía de un sistema como este.

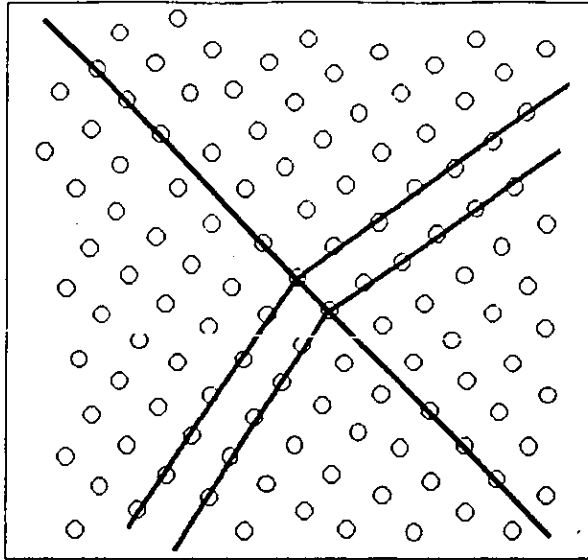


Figura 11-27 Detalle de la última configuración mostrada en la figura 11-26. Notese el eje de simetría en diagonal mostrado por la línea. Asimismo se muestran planos sobre los que se alinean los átomos a los lados del eje.

11.4 A futuro

- **Del programa** - El programa puede ser modificado en muchos aspectos. El más evidente es agregar más objetos que incrementen la funcionalidad y los algoritmos accesibles.

Es necesario modificar el programa para hacerlo más amigable con el usuario. Aún hay muchos aspectos que no son evidentes y tarde o temprano complican la vida. En particular el manejo de unidades debe ser modificado para permitir ingresar unidades reales y dejar que el programa haga las conversiones correspondientes.

Dado que las propuestas a futuro respecto a las fronteras de grano son de simulaciones de más larga duración y con más átomos es necesario estudiar el programa para mejorar su eficiencia. Sabemos que ahora el programa es versátil pero no muy eficiente, y que unas semanas de trabajo pueden mejorar su velocidad quizás hasta en más de un orden de magnitud.

La eficiencia del programa se puede mejorar utilizando radios de corte óptimos, para los que la conservación de la energía sea máxima y la cantidad de vecinos mínima.

Para tener representaciones más claras de lo que se ha observado podemos modificar los programas que muestran los átomos para que muestren vectores que representen los desplazamientos atómicos entre las configuraciones inicial y final. De esta forma será más clara la observación de los cambios.

- **De las fronteras de grano** - Analizados las simulaciones hay que aceptar que hace falta aún mucho trabajo.

Es necesario repetir las simulaciones para muchas otras fronteras, no sólo las especiales comparables con otros modelos. sino aprovechar que el modelo GCSN puede describir muchas otras.

Tanto para las fronteras que no simulamos como para las que si simulamos es necesario hacer simulaciones de mucho mayor duración, permitiendo equilibrios más confiables. pero sobre todo disminuyendo la velocidad de enfriamiento para garantizar un mejor mínimo de energía. En los casos de enfriamiento hay que enfriar de manera constante en cuanto a la disminución de la temperatura, de manera que puedan ser comparadas con los experimentos.

Las simulaciones a temperatura constante también tienen que ser de mucho mayor duración para poder garantizar la estabilidad y medir con más precisión la temperatura de recristalización.

Vale la pena estudiar con cuidado la aparición de las nuevas simetrías durante la simulación, observando las circunstancias en las que aparecen y midiendo sus propiedades. No sabemos si se presentan en todas las fronteras.

Utilizar mayor número de planos atómicos puede ser necesario para reducir la interacción entre una frontera GCSN y una de la *competencia*. También es posible eliminar esta interacción modificando las condiciones de frontera y utilizar una combinación de periódicas con fijas o puras fijas.

Fronteras fijas tiene la ventaja adicional de permitir modificar el volumen (densidad), para poder observar recristalización.

Hay diferentes variantes posibles dentro del modelo GCSN, personalmente me parece interesante modificar el modelo para suponer una frontera de 3 planos atómicos, y no de uno, en el que los dos planos laterales se construyen como dicta el modelo original, y el plano central es una iteración del proceso.

Parte V

Apéndices

A. HARDWARE Y SOFTWARE

A.1 La computadora

“*Creo que hay mercado en el mundo para unas cinco computadoras*”, Thomas Watson, fundador de IBM [Dic99].

Las necesidades de cómputo pueden variar enormemente de un programa de simulación a otro, por ejemplo. un programa que simule movimientos de un sistema planetario requiere de poca memoria para almacenar los planetas, pero con máxima precisión al calcular las interacciones y sus correcciones relativistas; un programa que simule un gas de esferas duras requiere de mucha memoria para las partículas, pero las interacciones son relativamente simples. Según [Erc97] comparada con otras aplicaciones, de la física computacional de la materia condensada. las simulaciones de dinámica molecular que usan potenciales clásicos no son particularmente demandantes, salvo que uno este interesado en tratar casos extremos. Prueba de esto es el hecho de que pudimos realizar los cálculos de esta tesis en computadoras personales.

El tiempo que requiere una computadora para llevar a cabo una simulación puede variar drásticamente de una máquina a otra, porque depende de una infinidad de factores como la velocidad de la memoria, la velocidad de almacenamiento en disco, el algoritmo usado, la velocidad de cómputo, otros trabajos pendientes, etc.

Además de los cálculos de estructuras electrónicas y de Monte Carlo, quizás el tema que más fuertemente compite por tiempo de cómputo con la dinámica molecular es el de predicción climática, por diversas razones:

- La *dinámica molecular* *no* tiene “prisa”, es decir, puede esperarse a encontrar los resultados, que, salvo por la competencia profesional no son urgentes. Como las simulaciones de dinámica molecular generan una enorme cantidad de resultados a analizar, tiempos de ejecución de una semana no necesariamente es malo porque el tiempo puede ser aprovechado para el análisis y la interpretación.
- La *predicción climática* *si* tiene “prisa”. Es muy compleja, normalmente con complejas condiciones iniciales. Estos datos pueden ser recopilados por diferentes métodos e introducidos en una computadora el mismo día con la intención de obtener predicciones climáticas. a lo más, algunos días después. Entre mayor sea la capacidad de cálculo antes se obtendrán los resultados y/o mejores predicciones se podrán hacer. Salvo por el interés académico, no nos sirve introducir condiciones del lunes, para predecir el clima del miércoles, ¡si los resultados no estarán listos hasta el viernes!.

A.1.1 ¿Que computadora?

Tratar de seleccionar *la mejor computadora para dinámica molecular* no es fácil. La elección esta condicionada por diversos factores:

- El tamaño del problema físico a estudiar.
- Otros usos que se le puedan dar a la máquina.
- ¿Se usara para generar gráficos?, ¿que tanto?. ¿de que calidad? y otras preguntas relacionadas con la visualización.
- Lo amigable del sistema operativo. el compilador y las herramientas de desarrollo a las que tengamos acceso en esa máquina (incluyendo previas experiencias y las características propias del sistema)
- Soporte técnico accesible y eficiente.
- Existencia de máquinas similares al alcance. disminuyendo los problemas de mantenimiento.

- Gente accesible y experimentada con la que podamos compartir opiniones y obtener apoyo en casos de necesidad.
- Precio.

Existen muchos, muchos factores más que pueden influenciar la compra o elección de una máquina, y dependiendo de la situación específica se dan los pesos adecuados a cada factor.

A.1.2 Almacenamiento

“No se puede tenerlo todo, ¿dónde lo pondrías?” Steven Wright.

Una regla conocida para calcular la cantidad de espacio de almacenamiento necesario en una computadora es: “Calcule el espacio requerido por los programas que le interesan, incluyendo sistema operativo, aplicaciones, utilerías y programas propios. Multiplíquelo por dos”, sin embargo también hay una máxima computacional reconocida: Nunca se tiene suficiente espacio de almacenamiento. Esto significa diseñar los programas considerando las limitaciones de la máquina en uso.

En el caso de los programas de esta tesis encontramos que almacenar las posiciones, velocidades y aceleraciones de 10,000 átomos con una precisión de 64-bits requiere de 720kb. También se requiere espacio para almacenar las tablas con los valores de las funciones 8.5, por ejemplo, para almacenar una tabla de un potencial de pares desde $r_{\min} = .5$ hasta $r_{\max} = 6.0$, con una precisión de 0.001 y usando 64-bits por dato (8 bytes) se requiere:

$$8 \text{ bytes} \times 5500 = 44000 \text{ bytes} \simeq 43 \text{ kb} \quad (\text{A.1})$$

que es muy poco para las máquinas modernas con varios Megabytes (10^6 kb) de memoria RAM y algunos Gigabytes (10^9 kb) de espacio en disco duro. Este estimado es conservador y debe revisarse la validez. Un programa aparentemente seguro en memoria puede convertirse fácilmente en una trampa. Por ejemplo, si en el programa anterior comenzamos a utilizar cuadruple precisión en los cálculos (128-bit) se doblaría inmediatamente el espacio requerido. Aumentar un orden de magnitud la precisión de las tablas lo aumenta también en un orden de magnitud. No sería raro tampoco, duplicar el alcance de las tablas, sobre todo en potenciales de tipo EAM 7.6.1. En este tipo de potenciales son necesarias varias tablas que contengan las densidades y sus derivadas. En este caso habríamos requerido 120 veces más espacio, unos 5160 kb \simeq 5 megabytes. Ya no tan despreciable.

En los últimos años ha aparecido un nuevo tipo de memoria, la memoria ultra rápida, generalmente limitada a unos cuantos kb (256 ó 512 es lo usual en computadoras personales), y reservada para la información más utilizada. Algunos lenguajes de programación permiten especificar las variables que queremos almacenar en este tipo de memoria bajo el nombre de variables de registro. Haciendo mediciones del programa de esta tesis encontramos que se puede reducir hasta en un 30% el tiempo de cálculo. Los compiladores más modernos (como el Visual C++ 6.0 que usamos) cuentan con algoritmos complejos para elegir las variables de registro y ya no es necesario preocuparnos por ellas. Las diferencias de velocidad en algunos de los procesadores personales más nuevos, como el Pentium III de Intel o el K6-3 de AMD se deben, en parte, a la adición de más memoria para contener variables de registro.

A.1.3 El CPU

CPU es el acrónimo que recibe el “corazón” una máquina. Se debe al inglés de *Central Processing Unit* o unidad de procesamiento central. Es la parte de la máquina encargada de procesar las instrucciones y entregar resultados.

El CPU tiene construido en sus circuitos una serie de procedimientos, altamente eficientes, para ejecutar algunas instrucciones, y únicamente son capaces de ejecutar instrucciones pertenecientes a este conjunto. Estas son relativamente simples y están fuertemente relacionadas con las características del aparato, por ejemplo, “carga al registro A el contenido de la localización de memoria 00900752”, o “multiplica los números contenidos en los registros A y B, almacenando el resultado en el registro C” son dos ejemplos de instrucciones a nivel máquina. Un programa escrito en lenguaje de máquina (conocido como *lenguaje ensamblador*) no es fácilmente transferible a otra computadora. Es muy difícil de decifrar para los humanos, porque su intención es ser muy fácilmente entendibles por la máquina. Originalmente si se buscaba aprovechar al máximo las capacidades de una máquina era necesario programar de esta forma.

Existe un debate entre dos diferentes maneras de diseñar los procesadores de una máquina. Un grupo propone construirlos con un mínimo de instrucciones de máxima eficiencia y reducir las instrucciones complejas a combinaciones de las básicas, optimizando el tiempo de ejecución (inclusive al nivel de requerir solamente un ciclo de reloj por instrucción ejecutada). Esta filosofía da lugar a los procesadores RISC (del inglés para *Reduced Instruction Set Computer*). Otro grupo propone una arquitectura que maneje instrucciones complejas, que requieran mayor tiempo de ejecución, pero que no requieran de tantos niveles de descomposición.

Una manera de describir la velocidad de una máquina es con el número de ciclos de reloj que lleva a cabo por segundo (medido en hertz). La máquina puede realizar un cierto número de operaciones básicas por ciclo de reloj, lo que a su vez determina la cantidad de operaciones básicas por segundo, y aún cuando esta puede ser de millones de operaciones por segundo, no significa que podamos mover millones de átomos por segundo, porque las operaciones a las que nos referimos son muy sencillas, por lo general solamente sumas, multiplicaciones y operaciones de este tipo.

Sea cual sea la máquina, el rendimiento será seguramente el punto más importante y la forma más común de evaluarlo es utilizar programas de prueba o de referencia (*benchmark* en inglés): programas que lleven a cabo una tarea específica (en ocasiones muy simple) en condiciones específicas y bien controladas. Al ejecutar un programa de prueba para evaluar una máquina se mide el tiempo de ejecución, y el resultado se compara con una referencia.

Existen muchos programas de prueba científicos, unos conocidos son el Linpack¹ y el SPEC². También se pueden usar programas no diseñados específicamente para medir capacidades científicas, aunque es más complicado evaluar los resultados y extender su validez a la dinámica molecular.

El problema es que diferentes máquinas se comportan de diferentes maneras con diferentes programas. En otras palabras, la *relación* entre el tiempo de ejecución entre dos programas de prueba puede variar notablemente de una máquina a otra. Linpack, por ejemplo, es un programa de prueba de álgebra lineal, de tal manera que máquinas que sean capaces de realizar multiplicaciones de matrices en forma eficiente (por una buena organización de su subsistema de punto flotante, compiladores altamente optimizados en este rubro, manejo de memoria u otros) obtendrán una buena puntuación en Linpack. A pesar de representar una clase extremadamente importante de problemas físicos (problemas de diagonalización por ejemplo), estas capacidades serán de muy poca utilidad para probar el desempeño en dinámica molecular porque no involucra problemas de álgebra lineal.

El otro extremo es ejecutar uno mismo una copia del programa que nos interesa en la máquina que nos interesa. Esto tiene la ventaja de que nos da una idea muy precisa de las capacidades de esa máquina en el problema específico que nos interesa. No siempre es una solución posible y práctica, ya sea porque la máquina no es accesible tan fácilmente, porque los tiempos requeridos para instalar y ejecutar el programa son demasiado largos, porque el programa aún no está disponible o alguna otra. Hoy en día se agudiza el problema de tener una máquina disponible, ya que la gran mayoría de las máquinas son configuradas según los requerimientos particulares de un cliente, con diferentes opciones que pueden modificar drásticamente el desempeño. Estas máquinas normalmente son ensambladas ya pagadas y muy poco antes de ser entregadas y probarlas es imposible.

Una opción es utilizar el programa de prueba MDBENCH³, específico para dinámica molecular. El autor ha compilado una buena cantidad de tiempos de ejecución en diferentes máquinas (uno mismo puede contribuir sus resultados). A pesar de que este programa puede no probar las propiedades específicas que cada uno de nosotros puede necesitar, es un buen indicador del desempeño en dinámica molecular.

A.2 Lenguaje

“Unix no fué diseñado para evitar que la gente haga cosas estupidas porque también evitaría que hiciera cosas inteligentes.”, Doug Gwyn [Han99].

El lenguaje de computadora es la interface entre el humano y la máquina. En la figura 62A-1 muestro los diferentes niveles de algunos lenguajes de programación. La traducción entre el humano y la máquina puede llevarse a cabo en dos momentos diferentes: i) Pueden traducirse las instrucciones conforme se vaya necesitando. ii) Pueden traducirse en algún momento entre que se escriben y se necesitan, en una fase llamada *compilación*.

Una vez que se ha escrito un programa en algún lenguaje, este debe ser convertido en las instrucciones de la máquina, este proceso puede hacerse antes de ejecutar el programa o al momento.

¹<http://uctlib.bell-labs.com/>

²<http://www.specbench.org/>

³<http://www.sissa.it/furio/mdbench.html>

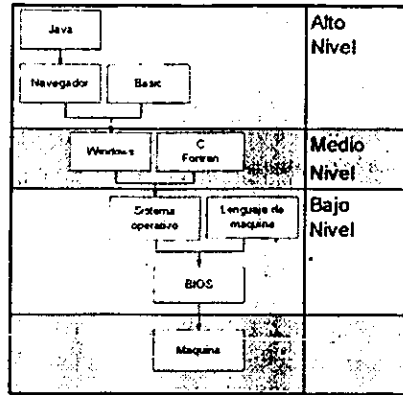


Figura A-1 Diagrama con los diferentes niveles de los lenguajes de programación, dependiendo de su cercanía a la máquina.

Hacer la traducción al momento tiene la ventaja de solo traducir lo que se va necesitando, pero la traducción requiere de mucho tiempo de cálculo y puede retrasar notablemente la ejecución. Si un programa repite los mismos pasos una y otra vez es evidente la poca eficiencia de traducirlos al momento, por lo que se desarrollaron lenguajes compiladores.

Durante la compilación, el programa compilador estudia el código y lo desmenuza en instrucciones básicas lo más eficientes posible. Al terminar genera un programa *ejecutable*. El ejecutable esta escrito ya en lenguaje de máquina, y si el compilador es bueno, puede ser muy eficiente.

Utilizar lenguajes diferentes del de máquina implica siempre un sacrificio en la posibilidad de controlar el 100% de una máquina, pero programar en lenguaje ensamblador es tan complicado que de cualquier manera esto es practicamente inconseguible.

Utilizar lenguajes diferentes al de máquina permite diseñar estándares para todas las máquinas, de tal manera que todas se programen de la misma manera, y que las diferencias sean adecuadas por el compilador al traducir las instrucciones al lenguaje específico de cada máquina. Esto permite a programadores de diferentes máquinas trabajar en conjunto.

Los lenguajes de programación independientes de la máquina se convierten entonces no sólo en una interface entre la máquina y los humanos, sino en un *lenguaje de comunicación entre humanos*. Un programa lleva con él la información de *cómo* se esta resolviendo un problema, y es extremadamente importante presentar esta información en una forma completa y clara para los que examien el procedimiento. Un lenguaje de programación no puede ser ambiguo. El código mismo es a la vez la mejor documentación de un programa y es muy importante tener en mente las necesidades de las personas (y no sólo de las maquinas) al escribirlo. En la práctica esto significa escribir comentarios informativos en el código, usar nombres claros e informativos para las variables, evitando los "trucos sucios" siempre que no sean absolutamente necesarios. Agregar estas capacidades de auto-explicación a un programa debe hacerse siempre al mismo tiempo que se escriba el código, y no dejarlo para después, cuando ya no necesariamente sea igualmente claro.

Los lenguajes se clasifican en dos grupos:

1. Lenguajes de Bajo Nivel - Las instrucciones en estos lenguajes son más cercanas a nosotros que las del lenguaje de máquina, pero se mantienen relativamente simples y específicas. Por lo general una instrucción en un lenguaje de este tipo se traduce en unas cuantas instrucciones básicas. Ejemplos son el Fortran y el C. Por lo general son lenguajes compiladores, y logran un alto grado de eficiencia, además de que permiten controlar casi por completo una máquina.

Hay quienes consideran al Fortran 77 un lenguaje de alto nivel, y al C de medio nivel, reservando el bajo nivel para el ensamblador

1. Lenguajes de Alto Nivel - Sus instrucciones son mucho más cercanas a nosotros, incluso que las de los lenguajes de bajo nivel. Por lo general una instrucción representa una gran cantidad de instrucciones básicas. Ejemplo es el Basic. Por lo general no son tan eficientes como los de bajo nivel, ni permiten el control total de la máquina, pero son más fáciles de aprender.

A.2.1 Basic

“Es prácticamente imposible enseñar un buen estilo de programación a estudiantes que han tenido una exposición previa a BASIC. Como programadores potenciales han sido mutilados más allá de lo recuperable”, E. W. Dijkstra [Han99].

En el basic las instrucciones, por lo general, son simples y poderosas. englobando una enorme cantidad de acciones de la máquina en una pequeña y clara instrucción. Esta simplicidad acarrea un problema: la poca versatilidad. Dado que las instrucciones de basic estan pensadas para ser simples y poderosas no en todas ellas es posible elegir entre la gran cantidad de parámetros disponibles en la máquina. Cuando los parámetros no son facilmente accesibles existe un valor por defecto (el que se supone sera util para la mayor parte de la población). Si con basic no podemos pedirle a la máquina que haga lo que queremos, podemos: i) Ampliar las instrucciones de basic. ii) Utilizar otro lenguaje más poderoso.

Basic es un lenguaje de ejecución, es decir, la traducción se lleva a cabo al momento de ejecutarse, lo cual lo hace demasiado lento para las aplicaciones numéricas. Recientemente han aparecido nuevas versiones de lenguajes Basic que permiten compilar los programas, mejorando notablemente el desempeño.

La más reciente versión de Basic, el Microsoft Visual Basic 6.0 es simple y muy poderoso. Simplifica notablemente la programación para ambiente windows y es una demostración de que Basic esta muy lejos de morir.

A.2.2 ¿Fortran o C?

“El objetivo principal del comando DATA es dar nombres a constantes; en lugar de referirse a pi como 3.141592653589793 en cada aparición, la variable pi puede ser dada ese valor con un comando DATA y usada en lugar de la forma más larga de la constante. Esto también simplifica modificar un programa en caso de que el valor de pi cambie.”, Manual de Fortran para computadoras Xerox [Han99].

Fortran viene de de las palabras formula translator, y desde su concepción hace más de treinta años (a finales de los 50's y desarrollado en los 60's y 70's) las versiones reciben el nombre correspondiente al lenguaje junto con el año de su estandarización (Fortran 77 será entonces el estándar de 1977 y Fortran 90 será el estándar de 1990). Para muchos científicos Fortran 77 ha sido el principal lenguaje computacional, y durante este tiempo sus capacidades numéricas son consideradas por muchos como superiores y más estables que las de ningún otro lenguaje de computadora. Los principales cambios han sido aumentar la diversidad y confiabilidad de las librerías numéricas. La unión de Fortran, técnicas para su uso y una gran cantidad de librerías numéricas constituyen la infraestructura dominante en la ciencia computacional.

C es un lenguaje mucho más reciente, desarrollado en los 70's provee las bases del sistema operativo Unix. La popularidad de Unix es una consecuencia directa del hecho de estar desarrollado en gran medida en C, y por lo tanto es facilmente portátil entre diferentes computadoras. C es un lenguaje simple y muy poderoso, con objetivos muy amplios. Problemas de ámbitos muy diferentes pueden ser solucionados con C efectivamente.

Los lenguajes de computadora evolucionan mucho más despacio que las computadoras mismas. La principal razón de esto es que trasladar programas a la nueva máquina muchas veces significa recompilarlos.

En la última década, sin embargo, la creciente importancia de estructuras de datos dinámicas (particularmente de arreglos dinámicos), estaciones de trabajo con Unix, sofisticadas capacidades de visualización interactiva y, mas recientemente, arquitecturas paralelas han hecho evidentes algunas incapacidades de Fortran 77, encendiendo el interés en el uso de otros lenguajes, sobre todo C. Recientemente C++ ha generado también considerable interés y Fortran ha tratado de trabajar sobre sus deficiencias evolucionando en el Fortran 90. Según el libro electrónico del Computational Science Education Project ([Cse95a]) una comparación entre los principales lenguajes científicos puede resumirse en la tabla A.1. en donde se incluyen los dos sabores de C (C y C++) y de Fortran (77 y 90). En la tabla una puntuación de 1 significa máxima capacidad y una de 4 significa mínima capacidad. La tabla A.1 representa el punto de vista del CSEP, no el mío. lo incluyo aquí para dar una perspectiva más global del problema y mostrar que aún cuando nosotros elegimos C para esta tesis, la posibilidad de elegir Fortran no es tan lejana.

Una de las principales ventajas de Fortran sobre C es su tradición en la física computacional, por lo que existen no sólo una gran cantidad de libros y artículos sobre algoritmos y sus propiedades, sino una decidida intención de formar gente con la habilidad de usar Fortran. Es difícil modificar una tradición de muchos años sobre todo sin razones decisivas. Las opciones completas son Fortran 90 y C++, mientras que C y Fortran 77 se consideran más bien herramientas limitadas algo obsoletas.

Funcionalidad	F77	C	C++	F90
Robusto numericamente	2	4	3	1
Paralelismo	3	3	3	1
Abstracción de datos	4	3	2	1
Programación orientada a Objetos	4	3	1	2
Programación funcional	4	3	2	1
Promedio	3.4	3.2	2.2	1.2

Tabla A.1 Comparación entre diferentes lenguajes y sus versiones más populares (Erc97).

Si la tabla A.1 representa el punto de vista de la CSEP y no la mía, me parece que sería conveniente comentar algunas de las ventajas en favor de cada uno de los lenguajes principales, alimentando una elección informada.

En favor de C:

- Existen compiladores eficientes para todas las arquitecturas en uso, además de un buen compilador público (gcc). Los compiladores de C son comúnmente incluidos gratis con las máquinas, mientras que los de Fortran 90 deben adquirirse por separado y pueden llegar a ser muy caros.
- C es muy amplio en espectro, y es mucho más poderoso que Fortran 90 en algunas áreas, sobre todo en el manejo de apuntadores y la manipulación de cadenas de caracteres.
- La experiencia adquirida para programar con C puede aplicarse directamente fuera del mundo científico, y C es el lenguaje más común en el mundo comercial.
- Los ambientes de programación para C tienen muchas veces más usuarios que los ambientes para Fortran 90, por lo que algunas rutinas de compilación han sido más trabajadas y pueden ser más eficientes. Estos ambientes de trabajo pueden facilitar enormemente el trabajo de depurar un programa, algo que no ocurre en Fortran 90 porque su más reducida base de usuarios es además menos exigente en este sentido.

A favor de Fortran 90:

- Un lenguaje diseñado con la intención de facilitar la programación de problemas numéricos como principal objetivo (por ejemplo operaciones con arreglos).
- Un lenguaje diseñado con la intención de generar código altamente eficiente en aplicaciones numéricas.
- Las rutinas y opciones de compilación han sido optimizadas para aplicaciones numéricas y no para aplicaciones de Internet o multimedia como muchas de C++.
- Según Ercolessi en [Erc97] existe una mejor capacidad de pasar código entre diferentes arquitecturas (entre máquinas de 32 y de 64 bit por ejemplo).

Ambos lenguajes pueden utilizar código de las versiones anteriores y es necesario aprender todas las virtudes de la nueva versión para continuar desarrollando y mantener las aplicaciones viejas, sino que pueden irse extendiendo las opciones conforme son necesarias.

Cada quién puede elegir el lenguaje que mejor le parezca según el peso que de a cada uno de los puntos anteriores. De hecho la decisión puede ser tan cerrada que muchos de los científicos que se dedican hoy en día a la física computacional han aprendido ambos lenguajes y lo eligen según la situación específica.

Los sistemas operativos modernos como Unix y Windows permiten utilizar librerías precompiladas que pueden estar generadas en cualquiera de los lenguajes, por lo que es posible encontrar una mezcla de ambos en un mismo programa.

A.2.3 C++

La mayoría de los novatos tenemos dificultades en entender las diferencias entre C++ y C. Para ponerlo fácil: *C++ es una versión extendida y mejorada de C, diseñada para programación orientada a objetos*. C++ contiene y soporta por completo al lenguaje C, además de las extensiones orientadas a objetos, es decir, C es un subconjunto de C++. Dado que C++ está construido sobre C, no se puede programar en C++ sin entender C [Sch95].

Muchos programadores utilizan C, no C++, aunque la posibilidad de programar orientado a objetos es una virtud importante de C++, pero limitada a algunas aplicaciones.

En 1980 se agregaron a C las clases, la verificación del tipo y los argumentos de las funciones y su conversión, y algunos otros aspectos; el lenguaje se denominó "C con clases". En 1983/84 C con clases fue rediseñado, extendido y reimplantado; el lenguaje resultante fue denominado C++. Las extensiones más importantes fueron las funciones virtuales y la sobrecarga de operadores [Str94].

C++ tiene diversas propiedades que lo hacen particularmente útil para programar simulaciones de dinámica molecular, algunas de las que aprovechamos en esta tesis son las siguientes:

- La posibilidad de encapsular funciones, disminuyendo la interferencia entre diferentes partes del programa, limitando la posibilidad de dañar partes del código que trabajan bien al modificar otras.
- Las funciones virtuales, que permiten programar y ejecutar de forma mucho más eficiente. En C eran necesarios programas mucho más largos (y complicados) o menos eficientes, para conseguir los mismos resultados.
- Aprovechar librerías ya hechas, no sólo por Microsoft o Borland, sino por los mismos investigadores del Instituto de Física, con quienes pudimos intercambiar funciones fácilmente.
- Mucho más fácil extender y dar mantenimiento al programa.
- Código más elegante y fácil de leer.

El ambiente de programación que utilizamos es el Visual C++ 6.0, optimizado para C++ [Kru97] [Gur97], y con una gran cantidad de ayudas para diseñar y mantener los programas. La única razón para desperdiciar las oportunidades que brinda VC++ 6.0 es que queremos mantener la portabilidad (sección 8.11) y usar el programa en las Silicon Graphics.

Una razón suficiente para usar C++ en lugar de C o de cualquier versión de Fortran es la gran calidad de los compiladores accesibles para Windows. Tanto el Borland C++ Builder como el Visual C++ de Microsoft, son excelentes ambientes de trabajo, y aunque hay que aceptar que están limitados a generar programas para máquinas PC (DOS, 3.x, 95 y NT) son fáciles de usar y dan muy buenos resultados.

El nombre de C++ proviene del "operador incremento" de C. Este operador equivale a incrementar en una unidad el valor de una variable. De esta forma el significado de C++ es "C aumentado".

A.2.4 De Pascal y Java.

Pascal es un lenguaje poderoso más parecido a C que a Fortran o a Basic. Las principales virtudes de Pascal las comparte con C, pero no comparte muchas de las virtudes de C++. Lo más recomendable para alguien que programe en Pascal y que este interesado en hacer programación científica, es que migre a un lenguaje más moderno como C++ o Fortran 90, de esa forma podrá compartir experiencias y librerías con otros programadores.

Java surge como una opción para escribir programas que puedan ser ejecutados en cualquier plataforma. La necesidad proviene del crecimiento de Internet y la búsqueda de una opción para difundir programas masivamente.

Los programas de Java no se traducen al momento como en el BASIC, sino que se compilan con anterioridad como en los lenguajes de alto nivel, pero con la diferencia de que no se realiza una traducción al lenguaje de máquina, sino que se traduce a un *lenguaje de máquina java* estándar. Cuando se ejecuta un programa en java en una máquina esta da instrucciones a la *maquina java*, esta a su vez traduce los comandos a la máquina específica. Con esto ganamos portabilidad porque los programas (y su compilación) son idénticos en cualquier máquina y pueden difundirse en Internet, pero perdemos eficiencia porque las instrucciones deben ser traducidas al momento por la máquina Java.

Java es recomendable para aplicaciones numéricas o cuando el programa esta pensado para ser ejecutado en máquinas iguales. Por esto la llamada "Revolución de Java" no pasa de ser una fiesta importante, pero no una revolución.

Programa	Uso	URL
Visual C++ 6.0	Compilador C++ para Windows.	www.microsoft.com
Visual Basic 6.0	Basic para Windows.	www.microsoft.com
ReView	Visor y animador en Windows.	www.brunel.ac.uk/depts/chem/ch241s/re_view/
Web Lab Viewer	Edito, visor y animador en Windows.	www.msi.com
Ras Mol	Visualización en Windows y Xwindows.	www.umass.edu/microbio/rasmol/
Fortran Workstation	Compilador Fortran 77 para Windows.	www.microsoft.com
Visual Diff	Comparador de archivos.	ninguno conocido - shareware
GWD text editor	Editor de texto que reconoce formatos.	www.iridis.com/gwd
fgraph	Graficador de datos.	ninguno conocido - shareware
Babel	Traductor de formatos.	ftp://joplin.biosci.arizona.edu/pub/Babel/babel16
Nail It	Fijar una pantalla como siempre visible.	www.common.net/~cv/
PCMI2D	Editor de configuraciones.	ninguno conocido - shareware
Periodic Table	Tabla periódica.	ninguno conocido - shareware
PerLib 32	Tabla periódica.	http://www.orbit.org/perlib/
Gnuplot y Wgnuplot	Graficador de datos de varias plataformas.	http://www.cs.dartmouth.edu/gnuplot
Scientific Workplace 2.5	Editor de textos científicos.	www.tcisoft.com/tcisoft.html
Adobe Acrobat	Visualizador de archivos PDF.	http://www.adobe.com/acrobat/
GsView	Visualizador de archivos PS.	www.cs.wisc.edu/~ghost/
Chem Sketch	Editor de configuraciones.	ninguno conocido - freeware
Winzip	Compresor de archivos.	www.winzip.com
WMF2EPS	Traductor de formato de gráficos.	schulter@temic.dofn.de
Mayura Draw	Editor y traductor de dibujos EPS.	www.mayura.com
PDQ-Lite	Editor de diagramas de flujo.	ninguno conocido - shareware

Tabla A.2 Programas útiles al programar simulaciones de dinámica molecular.

Sitio	URL
Base de datos de moléculas.	www.sci.ouc.bc.ca/chem/molecule/molecule.html
Biblioteca Molecular.	www.nyu.edu/pages/mathmol/library/library.html
Sitio de [Rap95].	www.cup.cam.ac.uk/onlinepubs/ArtMolecular/ArtMoleculartop.html
Dinámica Molecular.	xplor.csb.yale.edu/manual/htmlman/node205.html
Materials Society web page.	www.mrs.org/
Traductor Latex a HTML.	www-dsed.llnl.gov/files/programs/unix/latex2html/manual/
Molecular dynamics on the WWW.	www-dmse.mit.edu/courses/3.320/md.html

Tabla A.3 Sitios útiles al programar simulaciones de dinámica molecular.

A.3 Software existente

Durante el tiempo que trabaje en mi tesis y como ayudante en el curso de Física Computacional de la licenciatura en física de la UNAM junte diferentes programas y puntos de la red que pueden ayudar al que se dedique a la simulación en dinámica molecular, las tablas A.2 y A.3 resumen la información.

B. FORMATOS

B.1 Configuraciones

B.1.1 Formato PDB

El formato PDB obtiene su nombre del Protein Data Bank, y como es de esperarse esta optimizado para trabajar con moléculas orgánicas. Es muy completo en la información que incluye, aunque por su origen en aplicaciones de bioquímica, es normal sustituir algunos rubros del PDB inútiles fuera de la química orgánica y sustituirlos por otros adecuados. Podemos encontrarnos en situaciones extrañas al usar programas que esperan un PDB estándar y que no entiendan adecuadamente los cambios. Para una explicación de los rubros que sustituimos en este programa y su significado ir a la sección 9.9.1.

Los datos en el formato PDB tienen posiciones específicas, y no es posible modificarlas, por ejemplo, para definir posiciones atómicas con más precisión. El formato es muy estricto, y en caso de que no se requiera de alguno de los rubros, es necesario dejar los espacios en blanco. Es común que ocurra que un archivo PDB se vea correcto en un editor de texto, y que sin embargo, un programa lector de PDB's no pueda leerlo. La explicación puede ser que hayamos dado formato al archivo usando tabuladores o caracteres diferentes de espacio. En ese caso hay que revisar el archivo con un editor hexadecimal, que nos muestre el código ASCII de cada carácter.

El formato PDB es el siguiente:

Formato en C	Formato en Fortran	Columna	Descripción
char name[7]	A6	1-6	Tipo de registro (ej. ATOM, HETATM).
int id	I5	7-11	Número de átomo.
char Symbol[5]	A4	12-16 ¹	Nombre del átomo (ej. "CA", "ND1").
char coment[3]	A1	17	Posición alternativa (de ser necesaria).
char residual_name[5]	A3	18-21 ²	Código estándar de residuo de aminoácido.
char chainId[2]	A1	22	Cadena.
int resseq	I4	23-26	Número de secuencia de residuo.
char icode[2]	A1	27	Código de inserción (de ser necesaria)
		28-30	En blanco.
float x	F8.3	31-38	Coordenada X.
float y	F8.3	39-46	Coordenada Y.
float z	F8.3	47-54	Coordenada Z.
float occupancy	F6.2	55-60	Ocupación.
float tempfactor	F6.2	61-66	Factor térmico (valor-B).
		67	En blanco.
int fnote	I3	68-70	Número de nota.
char segId[5]		73-76	Código de segmento.
char element[3]		77-78	Elemento.
char charge[3]		79-80	Carga.
¹ es común dejar la columna 12 en blanco.		² es común dejar la columna 21 en blanco.	

Para leer archivos PDB sugiero utilizar en Silicon Graphics el Ras Mol, o el Cerius, y en Windows el Ras Mol. o el Web Lab Viewer (sección A.3).

B.1.2 Formato XYZ

Los archivos XYZ son más sencillos que los PDB, de hecho no he encontrado un formato XYZ estándar, y la mayoría de los programas que leen y escriben este formato son relativamente flexibles. Por lo general un archivo XYZ es tan simple como para incluir simplemente el número de átomo o el elemento que le corresponde y sus coordenadas en cada línea, sin necesitar definir las columnas específicas para cada dato.

La mayoría de los lectores y/o escritores de XYZ acostumbran definir al comienzo del archivo el número de átomos, y algunos pueden permitir una línea de comentario antes de la definición de las posiciones.

Los archivos XYZ recomiendo verlos en Silicon Graphics con el Ras Mol. o con el Cerius, y con el Ras Mol, el ReView o el Web Lab Viewer en Windows (sección A.3).

B.1.3 Para transformar de un formato a otro

Muchos programas que pueden leer y generar configuraciones sirven para transformar de un formato a otro. En Silicon Graphics recomiendo el Cerius y en Windows el Web Lab Viewer. Existe además un programa llamado Babel que puede transformar entre muchos formatos diferentes, existe para diferentes ambientes, y para Windows cuenta con una interface fácil de usar, aunque es un poco complicado de instalar (sección A.3).

B.2 Películas

Además de almacenar una configuración, podemos almacenar una serie de configuraciones, con la intención de animarlas y poder observar una película de la evolución.

Aunque existen películas que usan el formato PDB, la mayoría usan el formato XYZ, por su simplicidad, y sólo existen pequeñas variaciones de un programa de animación a otro.

Los programas que yo use y recomiendo son el AtomTV para Silicon Graphics y el ReView para Windows (sección A.3).

Los archivos para película son una serie de archivos XYZ comunes y corrientes, pero con la diferencia de que en AtomTV es necesario especificar al comienzo la cantidad de configuraciones que debe utilizar. Se puede dar un número menor del contenido por el archivo, pero uno mayor causa un error.

En ReView se incluye la línea de comentario después del número de átomos (que debe repetirse para cada configuración), aunque no es necesario numerar los átomos, basta con dar el elemento, lo que se usa para asignar un radio y color. ReView muestra en la pantalla de animación la línea de comentario, lo que es muy útil para seguir animaciones y entender a que situación corresponde una imagen.

Notese que si en ReView se intenta cargar una animación con demasiados átomos o configuraciones ocurre un error que cierra el programa. Cuando ReView lee un archivo detecta el elemento al que corresponde cada átomo, y automáticamente detecta las interacciones según una regla preestablecida. Si se supera un límite de 250 interacciones el programa se queja, pero no pasa nada grave. Es muy común tener demasiadas interacciones, sobre todo en archivos generados en unidades reducidas.

ReView está limitado a hacer animaciones en que la cantidad de átomos sea constante para todas las configuraciones. Atom TV no tiene esa limitación.

C. CONSTANTES Y FACTORES DE CONVERSIÓN

C.1 Factores de conversión

- $1.3806 \times 10^{-16} \text{ erg/átomo} = 1.987 \times 10^{-3} \text{ kcal/mol} = 8.314 \text{ J/mol}$
- $10^4 \text{ bar} = 1 \text{ GPa}$.
- $1 \text{ bar} = 10^5 \text{ Nm}^{-2}$ y $1 \text{ GPa} = 10^9 \text{ Nm}^{-2}$.
- Carga: $1 \text{ C} = 2.9979 \times 10^9 \text{ e.s.u.}$
- Dipolo: $1 \text{ C m} = 2.9979 \times 10^{11} \text{ e.s.u. cm}$
- Cuadrupolo: $1 \text{ C m}^2 = 2.9979 \times 10^{13} \text{ e.s.u. cm}^2$
- $1 \text{ cm} = 10^8 \text{ \AA}$
- $1 \text{ N} = 1 \text{ kg m/s}^2 = 10^5 \text{ g cm/s}^2 = 10^5 \text{ dyn}$
- $1 \text{ J} = 1 \text{ N m} = 10^7 \text{ dyn cm} = 10^7 \text{ ergs}$
- $1 \text{ bar} = 0.9869 \text{ atm} = 10^5 \text{ Pa} = 10^5 \text{ N/m}^2 = 10^5 \text{ J/m}^3$
- $1 \text{ N s/m}^2 = 10 \text{ g/(cm s)} = 10^3 \text{ cP}$

C.2 Constantes

- Número de Avogadro $N_A = 6.022 \times 10^{23} \text{ moléculas/mol}$
- Constante de Boltzmann $k_B = 1.381 \times 10^{-16} \text{ erg/(molécula } ^\circ\text{K)}$
- Constante de los gases $R = k_B N_A = 8.314 \text{ J/(Mol } ^\circ\text{K)} = 83.14 \text{ bars cm}^3/(\text{mol } ^\circ\text{K)}$
- Constante de Planck $h = 6.626 \times 10^{-27} \text{ erg s}$
- Permeabilidad eléctrica del vacío $\varepsilon_0 = 8.8542 \times 10^{-12} \text{ C}^2/(\text{N m}^2)$

C.2.1 Parámetros para el potencial de Lennard-Jones de argón

- $\sigma = 3.405 \text{ \AA}$
- $\varepsilon/k_B = 120^\circ\text{K}$

D. TALACHAS MATEMÁTICAS

D.1 Función del virial de Clausius

La medición de la presión en una simulación de dinámica molecular se basa en la función del virial de Clausius

$$\mathcal{W}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i^{total} \quad (D.1)$$

en donde \mathbf{f}_i^{total} es la fuerza total sobre el átomo i . Su promedio sobre el ensamble $\langle \mathcal{W} \rangle$ se obtiene, como el promedio sobre la trayectoria

$$\langle \mathcal{W} \rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t d\tau \sum_{i=1}^N \mathbf{r}_i \cdot m_i \ddot{\mathbf{r}}_i(\tau) \quad (D.2)$$

en donde ya hemos hecho uso de la segunda ley de Newton. Integrando por partes obtenemos

$$\langle \mathcal{W} \rangle = - \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t d\tau \sum_{i=1}^N m_i |\dot{\mathbf{r}}_i(\tau)|^2. \quad (D.3)$$

Esto es dos veces la energía cinética promedio, y por la ecuación 6.68 de equipartición

$$\langle \mathcal{W} \rangle = -DNk_B T \quad (D.4)$$

en donde D es la dimensionalidad del sistema, usualmente 3. N el número de partículas y k_B de la constante de Boltzmann.

Si pensamos en la fuerza total que actúa sobre una partícula de un sistema aislado como formada por dos contribuciones

$$\mathbf{f}_i^{total} = \mathbf{f}_i^{interna} + \mathbf{f}_i^{externa} \quad (D.5)$$

donde $\mathbf{f}_i^{interna}$ es la fuerza interna que surge de las interacciones interatómicas, y $\mathbf{f}_i^{externa}$ es la fuerza externa aplicada por las paredes del contenedor. Si las partículas se contienen en un paralelepípedo de lados L_x , L_y y L_z y volumen $V = L_x L_y L_z$, la parte $\langle \mathcal{W}^{externo} \rangle$ debido al contenedor puede evaluarse usando la definición de la ecuación D.1

$$\langle \mathcal{W}^{externo} \rangle = \mathcal{L}_x (-\mathcal{P} L_y L_z) + \mathcal{L}_y (-\mathcal{P} L_x L_z) + \mathcal{L}_z (-\mathcal{P} L_x L_y) = -DPV \quad (D.6)$$

donde, por ejemplo, $-\mathcal{P} L_y L_z$ es la fuerza externa $\mathbf{f}_x^{externa}$ aplicada por la pared yz sobre la dirección x a partículas localizadas en $x = L_x$. La ecuación D.4 puede entonces escribirse

$$\left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i \right\rangle - DPV = -DNk_B T \quad (D.7)$$

6

$$\mathcal{P}V = Nk_B T - \frac{1}{D} \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i \right\rangle \quad (D.8)$$

Este importante resultado se conoce con el nombre de *ecuación del virial*. Todas las cantidades, excepto la presión \mathcal{P} , son fácilmente accesibles en una simulación; ahora la ecuación D.8 nos provee de una manera para medir \mathcal{P} . La ecuación D.8 se reduce a la conocida ecuación de estado del gas ideal si las partículas no interaccionan unas con otras.

Para el caso de interacciones de pares por un potencial $\phi(r)$ la ecuación D.8 se transforma en

$$\mathcal{P}V = Nk_B T - \frac{1}{D} \left\langle \sum_i \sum_{j>i} r_{ij} \left. \frac{d\phi}{dr} \right|_{r_{ij}} \right\rangle \quad (D.9)$$

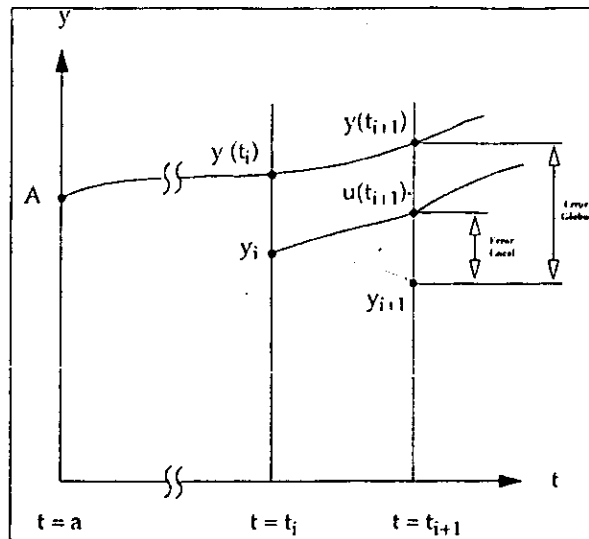


Figura D-1 Error local y error global en el paso t_{i+1} al integrar una ecuación diferencial. [Cse95a]

Esta expresión tiene la ventaja adicional sobre D.8 de estar naturalmente adaptada a ser usada con condiciones de frontera periódicas. Por supuesto no hay que olvidar que este método para obtener la ecuación del virial no es completamente válido, ya que si estamos suponiendo fronteras periódicas para el sistema, lo estaremos suponiendo infinito y no habrá fuerzas externas "reales". De cualquier manera el resultado es aplicable según lo mostrado por Erpenbeck y Wood en 1977 [Erp77].

D.2 Errores

"Una vida pasada cometiendo errores no sólo es más honorable, sino más útil que una pasada haciendo nada." George Bernard Shaw.

D.2.1 Errores de integración

Se acostumbra clasificar los errores en dos grupos: *errores globales* y *errores locales*:

- **Errores globales.** Son la diferencia que en un momento dado el sistema simulado tiene con un sistema real en el que no se hubiera hecho ninguna aproximación. Es el resultado de la simulación completa y por lo general es el error que nos interesa más conocer, además de ser el menos accesible.
- **Errores locales.** Es la diferencia entre el sistema real y el simulado acumulado en un intervalo de tiempo. La acumulación de los errores locales da lugar al error global.

La figura D-1 muestra la sutil diferencia entre ambos errores.

Por supuesto que el hablar de un sistema real tiene sus problemas filosóficos, pues no siempre es fácil definir a lo que nos referimos. Además existen los problemas cuánticos de definir la configuración y propiedades de un sistema, pero por un momento los dejo de lado porque se encuentran fuera del espectro de la tesis.

En el apéndice D.2.2 hago una derivación simple del error estadístico como función del tiempo de observación y el efecto de utilizar un sistema finito.

D.2.2 Error estadístico

Tanto en los métodos deterministas cómo en los estocásticos las configuraciones sucesivas están correlacionadas, ¿que significa esto si calculamos el promedio temporal de la observable \mathcal{A} , que necesariamente sólo puede incluir un tiempo finito?. Podemos considerar el error estadístico para n observaciones sucesivas $\mathcal{A}_1, \dots, \mathcal{A}_n$:

$$\langle (\delta\mathcal{A})^2 \rangle = \left\langle \left[\frac{1}{n} \sum_{i=1}^n (\mathcal{A}_i - \langle \mathcal{A} \rangle) \right]^2 \right\rangle \quad (\text{D.10})$$

Si recordamos la definición de la función de autocorrelación para la observable \mathcal{A}

$$\phi_{\mathcal{A}}(t) = \frac{\langle \mathcal{A}(0)\mathcal{A}(t) \rangle - \langle \mathcal{A} \rangle^2}{\langle \mathcal{A}^2 \rangle - \langle \mathcal{A} \rangle^2}. \quad (\text{D.11})$$

y del tiempo de correlación característico

$$\tau_{\mathcal{A}} = \int_0^{\infty} \phi_{\mathcal{A}}(t) dt \quad (\text{D.12})$$

podemos volver a escribir el error estadístico cómo

$$\langle (\delta\mathcal{A})^2 \rangle \simeq \frac{2\tau_{\mathcal{A}}}{n\delta t} \left[\langle \mathcal{A}^2 \rangle - \langle \mathcal{A} \rangle^2 \right] \quad (\text{D.13})$$

en donde δt es el tiempo entre observaciones (por lo que $n\delta t$ es el tiempo total de observación τ_{obs}).

Hay varias cosas que hacer notar de este error estadístico, lo primero es que no depende del tiempo entre las observaciones, sino del tiempo total de observación, de tal manera que si $n \rightarrow \infty$ entonces $\delta t \rightarrow 0$. También vemos que el error no es el que se encuentra si las observaciones fueran independientes. El error aumenta con el tiempo de correlación característico ($\tau_{\mathcal{A}}$) entre configuraciones, de tal manera que sólo una disminución en este tiempo y/o un aumento del número de observaciones (n) puede reducir el error.

Si el error estadístico para una observable \mathcal{A} depende del tiempo finito de observación, podemos preguntarnos por la dependencia en lo finito del sistema. Para ello definimos

$$\Delta(n, L) \equiv \sqrt{\frac{\left(\langle \mathcal{A}^2 \rangle_L - \langle \mathcal{A} \rangle_L^2 \right)}{n}} \quad (\text{D.14})$$

en donde la L se refiere a una dimensión lineal del sistema, y $\langle \cdot \rangle_L$ es un promedio con respecto al tamaño finito.

En el caso del equilibrio termodinámico para un sistema de tamaño infinito basta una observación para obtener \mathcal{A} , es decir, si $L \rightarrow \infty$ entonces $\Delta(n, L) \rightarrow 0$, independientemente de n : por otro lado, si incrementamos el tamaño del sistema, la cantidad efectiva de observaciones debe aumentar. Con L el tamaño del sistema y L' es el tamaño que se obtiene tras un escalamiento por un factor b con $b > 1$ ($L' = bL$). El número de observaciones efectivas cambia a $n' = b^{-d}n$ en donde d es el número de dimensiones.

Bibliografía

- [Aar72] Aaron H. B. y Bolling G.F., *Surface Science*, **31**, 27, (1972).
- [Ald57] Alder B. J., Wainwright J., *Phase transition for a hard sphere system*, J. Chem. Phys. **27**, 1208-1209 (1957).
- [Ali99] Alison's Favorite Quotations. Página del www: <http://tranquility.mit.edu/quotes/quotes.html>
- [All87] Allen M. P. y Tildesley D.J., *Computer simulation of liquids*, Oxford Science Publications, Gran Bretaña. (1987).
- [Ame99] Ames National laboratories, página de internet <http://cmp.ameslab.gov/cmp/cmp.html> mantenida por Morris J., Turner D. y Kai-Ming Ho.
- [And80] Andersen H. C., *Molecular dynamics simulations at constant pressure and/or temperature*, J. Chem. Phys., **72**, 2384-93, (1980).
- [And83] Andrea T. A. Swope W. C. y Andersen H. C., *The role of long ranged forces in determining the structure and properties of liquid water*, J. Chem. Phys., **79**, 4576-84, (1983).
- [And97] Andrews Mark, *Aprenda Visual C++ Ya*, Microsoft Press y McGraw-Hill, España, (1997).
- [Ara97] Aragón J. L., Romeu D., Beltrán del Río L., y Gómez A., *Acta Crystallogr., Sect. A: Found. Crystallogr.* **53**, 772, (1997).
- [Bal72] Balluffi R. W. y Schober T., *Scripta Met.*, **6**, 697, (1972).
- [Ban88] Banerjee A. y Smith J. R., *Origins of the universal binding-energy relation*, Phys. Rev. B, **37**, vol 12. 6632. (1988).
- [Ban92] Banks J., Brooks J., Cairns G., Davis G. y Stacy P., *On devaney's definition of chaos*, American Mathematical Monthly, **99**, 332-334, (1992).
- [Bar71] Barker J. A., Fisher R.A. y Watts R. O., *Liquid argon: Monte Carlo and molecular dynamics calculations*. Mol. Phys., **21**, 657-73, (1971).
- [Bas86] Baskes M. I., Foiles S. M. y Daw M. S.. Phys. Rev. B, **33**. (1986).
- [Bee76] Beeman D. J., J. Comp. Phys. **20**. 130. (1976).
- [Ben92] Bennet D. A., *Chaos and linearity*. Journal of undergraduate mathematics. **24**, 41-49, (1992).
- [Ber84] Berendsen H J C., Postma J. P. M., van Gunsteren W. F., DiNola A., Haak J. R., *Molecular dynamics with coupling to an external bath*, J. Chem. Phys., **81**, 3684-90. (1984).
- [Ber86] Berendsen H. J. C., y van Gunsteren W. F., en [Cic86], pág. 43.
- [Ber87] Berendsen H. J. C., pág 496 en [Cic86]. Comp. Phys. Commun. **44**, 23, (1969).
- [Bis68] Bishop G. H. y Chalmers B., *Scripta Met.*, **2**, 133, (1968).
- [Bol74] Bollman W. *Crystal defects and crystalline interfaces*, Heidelberg, Springer-Verlag, (1970).
- [Bol74] Bollman W., *Phys. Status Solidi*. **21**, 543-550. (1974).
- [Bol82] Bollman W., *Crystal Lattices. Interfaces. Matrices*. Bollman, Gèneve. (1982). Published by the author.

- [Boo72] Booth A. D., *Numerical methods*, 3rd edn., Butterworth, London, (1972).
- [Bor12] Born M. y von Karman Th., *Über schwingungen in Raumgittern*, Physick. Z., **12**, 297-309, (1912).
- [Bra40] Bragg W. L., Proc. Phys. Soc., **52**, 54, (1940).
- [Bra42] Bragg W. L., J. Sci. Instrum., **19**, 148-150, (1942).
- [Bra47] Bragg W. L. y Nye J. F., Proc. Roy. Soc., **A190**, 474-481, (1947).
- [Bra61] Brandon D. G y Wald M., *Phil. Mag*, **6**, 1035, (1961).
- [Bra64] Brandon D. G., Ralph B., Ranganathan S. y Wald M. S., *Acta Met.*, **12**, 813, (1964).
- [Bri98] Brillouin M., *Ann. Chem. Phys.*, **13**, 77, (1898).
- [Bri99] Bristol University Department of Physics Research Booklet. U.K. (1999).
- [Bro84] Brown D., Clark J. H. R., *A comparison of constant energy, constant temperature and constant pressure ensembles in molecular dynamics simulation of atomic liquids*, *Molecular Physics*, **51**, 1243-52, (1984).
- [Bro85] Brooks C. L. III, Mintgomery, Pettitt B. y Karplus M., *Structural and energetic effects of truncating long range interactions in ionic and polar fluids*, *J. Chem. Phys.*, **83m** 5897, 5908, (1985).
- [Bur40] Burgers J. M., Proc. Phys. Soc., **52**, 52, (1940).
- [Buz94] Buzzi-Ferraris Guido, *Scientific C++, Building numerical libraries the object-oriented way*, Addison-Wesley, (1994).
- [Car85] Car R. y Parrinello M., *Phys. Rev. Lett.*, **55**, 2471, (1985).
- [Car90] Carlsson A, *Solid State Phys.*, **43**, 1, (1990).
- [Car96] Carlberg M. H., *Molecular dynamics studies of metallic superlattices*, tesis de licenciatura N°574 Linköping Studies in Science and Technology, (1996).
- [Ceb98] Ceballos Francisco J., *Programación orientada a objetos con C++*, Alfaomega grupo editor, 2ª ed. (1998).
- [Cha61] Chapman S. y Cowling T. G., *The mathematical theory of non-uniform gases*, Cambridge University Press, London, págs. 26, 37 y 40-41, (1961).
- [Cha70] Chaudari O. y Matthews J. W., *Appl. Phys. Lett.*, **17**, 115, (1970).
- [Cha71] Chalmers B. y Gleiter H., *Phil. Mag.*, **24**, 1541, (1971).
- [Cha76] Chadwick G. A. y Smith D. A. (editores), *Grain boundary structure and properties*, Academic Press, Londres, (1976).
- [Che77] Cheung P. S. Y., *On the calculation of specific heats, thermal pressure coefficients, and compressibilities in molecular dynamics simulations*, *Mol. Phys.*, **33**, 519-26. (1977).
- [Chr88] Christman J. R., *Fundamentals of solid state physics*, John Wiley & Sons, EUA, (1988).
- [Cic86] Cicotti G. and Hoover W. G. (eds), *Molecular dynamics simulation of statistical-mechanical systems*, North-Holland (1986).
- [Cic87] Cicotti G., Frenkel D. y McDonald I.R. (eds), *Simulation of liquids and solids*. North-Holland (1987).
- [Cse95a] Computational Science Education Program, Libro electrónico: *Fortran 90 and computational science.*, <http://yellow.ccs.uky.edu/~ccs/csep/TEXTOC.html>.
- [Daw83] Daw M. S. y Baskes I., *Phys. Rev. Lett.*, **50**, 1285. (1983).
- [Daw84] Daw M. S. y Baskes M. L., *Phys. Rev. B*, **29**, 6443 (1984).
- [Daw89a] Daw M. S., *Phys. Rev. B*, **39**, 7441. (1989).

- [Daw89b] Daw M. S. en *Atomistic Simulation of Materials Beyond Pair Potential*, edited by V. Vitek y D. J. Srolovitz, Plenum, Nueva York (1989), p. 181.
- [Dev93] Devis Botella Ricardo, *Programación orientada a objetos en C++*, Paraninfo, Madrid, (1993).
- [Dev94] DeVries Paul L., *A first course in computational physics*, John Wiley & Sons Inc., USA, (1994).
- [Dev97] Devis Botella Ricardo, *C++, STL Plantillas Excepciones Roles y Objetos*. Paraninfo, (1997).
- [Dic99] Dictionary of Scientific Quotations, página de la www: <http://naturalscience.com/dsqhome.html>
- [Eas80] Eastwood J. W., Hockney R. W. y Lawrence D., *P3M3DP- the three dimensional periodic particle-particle/particle-mesh program*. Comput. Phys. Commun. **19**, 215-61, (1980).
- [Erc86] Ercolessi F., Tosatti E., Parrinello M., Surf. Sci., **177**, 314, (1986).
- [Erc88] Ercolessi F., Parinello M. y Tosatti E., Philos. Mag. A, **58**, 213, (1988).
- [Erc94] Ercolessi F. y Adams J. B., Europhys. Lett., **26**, 583, (1994).
- [Erc97] Ercolessi F., *A molecular dynamics premier*, Manual del International School for Advanced Studies en Trieste, Italia, 56 págs., (1987).
- [Erp77] Erpenbeck J. J. y Wood W. W., *Molecular dynamics techniques for hard core systems*, en *Statistical Mechanics B. Modern theoretical chemistry*. Vol. 6, 1-40, Plenum, Nueva York. Berne B. J. (Ed.), (1977).
- [Eva83a] Evans D. J., J. Chem. Phys., **78**, 3297. (1983).
- [Eva83b] Evans D. J., Hoover W. G., Failor B. H., Moran B. y Ladd A. J. C., Phys. Rev. A, **28**, 1016, (1983).
- [Eva84] Evans D. J. y Morris G. P., Comp. Phys. Rept. **1**, 297, (1984).
- [Fer98] Fernandez María Eufemia, *Estructura atómica de frontera de grano*, Tesis de licenciatura, Asesor Dr. David Romeu. IFUNAM, Mexico (1998).
- [Fey87] Feynman R. P. , Leighton R. B. y Sands M., *Física (Volumen I: Mecánica, radiación y calor)*, Addison-Wesley Iberoamericana, E.U.A., Oelker E. y Espinosa D. (trad.), (1987).
- [Fey97] Feyerabend P., *Tratado contra el método*. Ed. Rei, 1ª reimpresión, México, 1997.
- [Fin81] Fincham D. y Ralston B. J., *Molecular dynamics simulation using the CRAY-1 vector processing computer*. Comput. Phys. Commun., **23**, 127-34, (1981).
- [Fin84] Finnis M. W. y Sinclair J. E., Philos. Mag. A, **50**, 45, (1984).
- [Fin92] Finnis M. W. y Manfred R., *Structures of nterfaces in crystalline solids*, cap. 9 de [Ger92].
- [Fle80] Fletcher R., *Practical Methods of Optimization, Vol. 1, Unconstrained Optimization*, John Wiley & Sons. New York, (1980).
- [Foi86] Foiles S. M., Baskes M. I. y Daw M. S.. Phys. Rev. B, **33**. 7983. (1986).
- [For67] Fortes M. A. y Ralph B., *Acta Met.*, **15**. 707. (1967).
- [Fox84] Fox J. R. y Andersen H. C., *Molecular dynamics simulations of a supercooled monoatomic liquid and glass*. J. Phys. Chem., **88**, 4019-27, (1984).
- [Fre96] Frenkel Daan y Smit Berend, *Understanding molecular simulation, from algorithms to applications*, Academic Press, (1996).
- [Fri26] Friedel G., *Lecons de Cristallographie*. Paris. (1926).
- [Fuk55] Fukushima E. y Ookawa A.. J. Phys. Soc. Japan, **10**. 970-981. (1955).
- [Fuk57] Fukushima E. y Ookawa A.. J. Phys. Soc. Japan, **12**. 139-146. (1957).

- [Gar98] Garzón I. L., Kaplan I. G., Santamaria R. y Novaro O., *Molecular dynamics study of the Ag₆ cluster using ab initio many-body model potential*, Journal of chemical physics, **109**, 6, 2176, (1998).
- [Gas95] Gasiórowicz S., *Quantum physics*, John Wiley & Sons, 2nd Ed., (1995).
- [Gca71] Gear C. W., *Numerical initial value problems in ordinary differential equations*, Prentice Hall, E.U.A., (1971).
- [Ger92] Gerold V., *Structure of solids*, volumen 1 de *Materials Science and Technology. A comprehensive treatment*, editores Cahn R. W., Haasen P. y Kramer E. J., VCH, (1992).
- [Gib60] Gibson J. B., Goland A. N., Milgram M. y Vineyard G. H., *Dynamics of radiation damage*, Phys. Rev. **120** 1229-1253 (1960).
- [Gif67] Gifkins R. C., Mater. Sci. Eng., **2**, 181, (1967).
- [Gif69] Gifkins, E. C., *Interfaces conference proceedings, Melbourne, Australia, 1969-Review Papers*, p.29, The Australian Institute of Metals and Butterworths. (1969).
- [Gle72] Gleiter H. y Chalmers B., Prog. Mat. Sci., **15**, (1972).
- [Gol94] Goldstein H., *Mecánica Clásica*. Ed. Reverté, 2a Ed., Española, Fernández J. (trad.), (1994).
- [Gom98] Gómez A., Beltrán del Río L., Aragón J. L. y Romeu D., Scr. Mater., **38**, 795, (1998).
- [Gou96] Gould H., Tobochnik Jan, *An introduction to computer simulation methods*, Addison-Wesley, 2ªed., (1996).
- [Gur97] Gurewicz N. y Gurewicz O., *Teach yourself Visual C++ 5 in 21 days*, Sams Publishing, EUA, (1997).
- [Hag79] Hagler A. T., Stern P. S., Sharon R., Becker J. M. y Naider F., *Computer simulation of the conformational properties of oligopeptides. Comparison of theoretical methods and analysis of experimental results*, J. Am. Chem. Soc., **101**, 6842-6852, (1979).
- [Hai92] Haile J. M., *Molecular dynamics simulation, elementary methods*, John Wiley & Sons, EUA, (1992).
- [Han69] Hansen J. P. y Verlet L., Phys. Rev., **184**, 151. (1969).
- [Han86] Hansen J. P., McDonald I. R., *Theory of simple liquids*, 2nd Ed., Academic, (1986).
- [Han99] Hansen Stecen, *Quotes about computers and software and other things*, Página de la www: <http://members.tripod.com/~imarusic/compuquote.html>
- [Has71] Hasson G. y Goux C., *Scripta Met.*, **5**, 889, (1971).
- [Haw88] Hawking, Stephen W., *Historia del tiempo. Del big bang a los agujeros negros*, Ed. Grijalbo, México, (1988).
- [Hee82] Heermann D.W., Klein W., Stauffer D., Phys. Rev. Lett. **49**, 1262, (1982).
- [Hee90] Heermann D. W., *Computer simulation methods in theoretical physics*, Springer-Verlag, 2nd Ed., (1990).
- [Hey82] Heyes D.M. y Singer K., *A very accurate molecular dynamics algorithm*, CCP5 Quarterly, **6**, 11-23, (1982).
- [Hoc74] Hockney R. W., Goel S. P. y Eastwood J. W., J. Comp. Phys., **14**, 148, (1974).
- [Hof56] Hoffman R., *Acta Met.*, **4**, 97. (1956).
- [Hon72] Hondros E. D. y Seah M.P., *Scripta Met.*, **6**, 1007, (1972).
- [Hon76] Hondros E.D., McLean D., *Grain Boundary fragility*, cap. 10 de [Cha76], (1976).
- [Hoo85] Hooman S. W., *A molecular mechanical force field for the conformational analysis of oligosaccharides: Comparison of theoretical and crystal structures of Man alpha 1-3 Man beta 1-4 GlcNAc*, Biochemistry, **29**, 9110-9118, (1990).
- [Imu63] Inman W. C., McLean D. y Tipler H. R., *Proc. Roy. Soc.*, **A273**, 538. (1963).
- [Ish73] Ishida Y. McLean M. *Phil. Mag.*, **27**, 1125. (1973).

- [Ish76] Ishida Y., *The bubble raft as a model for grain boundary structure*, cap. 3 en [Cha76], pág 93-106, (1976).
- [Jac87] Jacobsen K. W., NØskov J.K. y Puska M. J., *Phys. Rev. B.* **35**, 7423, (1987).
- [Joh70] Johannesson T. y Thölen A., *Phil. Mag.*, **21**, 1223, (1970).
- [Joh88] Johnson R. A., *Phys. Rev. B.*, **37**, 3924, (1988).
- [Ke49] Ke T. S., *J. Appl. Phys.*, **20**, 274, (1949).
- [Ker91] Kernighan B. W. y Ritchie D. M., *El lenguaje de programación C*, 2a Ed., Prentice Hall, México. (1991).
- [Kir83] Kirkpatrick S., Gelatt C. D. y Vecchi M. P., *Science*, **220**, 671. (1983).
- [Kit86] Kittel C., *Introduction to solid state physics*, John Wiley & Sons, E.U.A, 6th Ed, (1986).
- [Knu68] Knuth D. E., *Seminumerical Algorithms*, vol. 2 de *The Art of Computer Programming*, Addison-Wesley, Reading, MA, (1973).
- [Koo90] Koonin Steven E. y Meredith Dawn C., *Computational Physics, Fortran Version*, Addison-Wesley, E.U.A., 635 págs, (1990).
- [Kro49] Kronberg M. L. y Wilson F. H., *Trans. A.I.M.E.*, **185**, 501. (1949).
- [Kru97] Kruglinski D. J., *Programación avanzada con Microsoft C++ Versión 5.0*. Microsoft Press-McGraw Hill, Córdoba R. y Domingo A. (trads.), México, (1997).
- [Lan94] Landau L. D. y Lifshitz E. M., *Mechanics (Course of theoretical physics: v.1)*, Pergamon, 3rd Ed., Gran Bretaña, Sykes J. B. y Bell J. S. (trad), (1976).
- [Lap51] Laplace de, P.S., *A Philosophical Essay on Probabilities* (transl.), Dover, Nueva York, (1951).
- [Lec70] Le Coze J., Ranson C. y Goux C., *C. R. Acad. Sc. Paris.*, **C271**, 1120, (1970).
- [Lob76] Loberg B. y Norden H., *High resolution microscopy of grain boundary structure*, cap. 1 en [Cha76]. (1976).
- [Lod99] Lodriguss, Jerry, *Catching the light. The art of deepsky astrophotography*, página de la [www: www.astrosurf.com/neptune/astropix/TOC.HTM](http://www.astrosurf.com/neptune/astropix/TOC.HTM)
- [Lom52] Lomer W. M. y Nye J. F., *Proc. Roy. Soc.*, **212**, 576-584. (1952).
- [Luc82] Luckhurst G. R. y Simpson P., *Computer simulation studies of anisotropic systems. VIII. The Lebwohl-Lasher model of nematogens revisited*, *Mol. Phys.*, **47**, 251-65. (1982).
- [Mac90] MacPherson Alistair K., *Atomic mechanics of solids*, 2º volumen de *Mechanics and physics of discrete systems*, editor George C. Sih., North-Holland. (1990).
- [Man76] Mandell M. J., *On the properties of a periodic fluid*, *J. Stat. Phys.* **15**, 299-305, (1976).
- [McI65] McLean y Mykura H., *Temperature Coefficient of Twin Boundary energy of Platinum and cobalt*, *Phil. Mag.*, **14**, 1191, (1965).
- [Met49] Metropolis N. y Ulam S., *J. Am. Stat. Ass.*, **44**. (1949).
- [Met53] Metropolis N., Rosenbluth A. W., Rosenbluth M. N. y Teller A. H., *J. Chem. Phys*, nr. 6, **21**, (1953).
- [Mic98] Microsoft Press (sin autor), *Visual Basic 6.0. Manual del programados*, Microsoft Press, (1998).
- [Mon94] Mondragón A., Lima C.D. y Wang J.C.. *Three-dimensional structure of the 67K N-terminal fragment of E coli DNA topoisomerase I*, *Nature* **367**. 133-146 (1994).
- [Mot48] Mott N. F., *Slip at grain boundary and grain growth in metals*. *Proc. Roy. Soc.*, **60**, 391, (1948).
- [Msi97] Molecular Simulation Inc., *The Discover Program Documentation*. [http://www.nchc.gov.tw/RESEARCH/Chem doc/discover/Disco_Home.html](http://www.nchc.gov.tw/RESEARCH/Chem_doc/discover/Disco_Home.html). (1997).

- [Mül97] Müller-Plathe Florian, *Molecular simulation - A Primer*, A ser publicado en: Wilson S. y Diercksen G.H.F (editores), *Problem solving in computational molecular science: Molecules in different environments*, NATO ASI Series B Physics Vol. ?, (Kluwer Academic Publishers. Dordrecht, 1997?).
- [Mül98] Müller-Plathe Florian, *Computational chemistry I. Molecular Simulation*, lecture notes collected by faller R., Hahn O., Mooney D., Reith D., Schmitz H. y Schmitz R., (1998)
- [Mün69] Münster A., *Statistical thermodynamics*. Springer/Academic Press, Berlin/New York, 1969.
- [Nic79] Nicolas J. J., Gubbins K. E., Strett W. B. y Tildesley D. J., *Equation of state for the Lennard-Jones fluid*, Mol. Phys., **37**, 1429-54, (1979).
- [Nos84a] Nosé S., *A molecular dynamics method for simulations in the canonical ensemble*, Molec. Phys., **52**, 255-268, (1984).
- [Nos84b] Nosé S., *A unified formulation of the constant temperature molecular dynamics methods*, J. Chem. Phys., **81**, 511-519, (1984).
- [Nos91] Nosé S., *Constant temperature molecular dynamics methods*, Prog. Theoret. Phys. Supplement, **103**, 1-46, (1991).
- [Par81] Parrinello M. y Rahman A., *Polymorphic transitions in single crystals: A new molecular dynamics method*, J. Appl. Phys., **52**, 7182-7190, (1981).
- [Pen79] Penrose O. y Lebowitz J. L. en *Studies in statistical mechanics*, Vol. 7, Lebowitz J. L. y Montroll E. W. (Eds.), North-Holland, Holanda, (1979).
- [Peñ91] Peña, L. de la, *Introducción a la mecánica cuántica*, Fondo de cultura económica, México, (1991).
- [Pon74] Pond R. C. y Smith D. A., *Can. Met. Quart.*, **13**, 39, (1974).
- [Pon76] Pond R. C. y Smith D. A., *Bollmandd's O-lattice theory: a geometrical approach to interface structure*, International Metals Reviews, June 1976, p. 61.
- [Pow82] Powles J. G., Evans W. A. B. y Quirke N., *Non-destructive molecular dynamics simulation of the chemical potential of a fluid*. Mol. Phys, **46**, 1347-70, (1982).
- [Pow84a] Powles J. G., *Singer on the root*, CCP5 Quarterly, **11**, 39-41, (1984).
- [Pow84b] Powles J. G., *The liuvuid-vapour coexistence line for Lennard-Jones-type fluids* Physics, **A126**, 289-99, (1984).
- [Pra81] Praat L. R. y Haan S. W., *Effects of periodic boundary conditions on equilibrium properties of computer simulated fluids. I. Theory*, J. Chem. Phys., **74**, 1864-72, (1981).
- [Pre86] Press W. H., Flannery B. P., Teukolsky S. A. y Vetterling W. T., En *Numerical recipes, The art of scientific computing*. Cambridge University Press, Cambridge, (1986).
- [Pun72] Punphrey P. H., *Scripta Met.*, **6**, 107, (1972).
- [Pun76] Punphrey P. H., *Special High Angle Grain Boundaries*, cap. 5 en [Cha76], (1976).
- [Qui05] Quincke G., *Proc. Roy. Soc.*, **A76**, 431, (1905).
- [Rah64] Rahman A., *Correlation in the motion of atoms in liquid argon*, Phys. Rev. **136** A405-A411.
- [Ran65] Ranganathan S., Ph. D. Thesis, Cambridge, (1965).
- [Ran66] Ranganathan S., *Acta Crystallogr.*, **21**, 197, (1966).
- [Rap80] Rapaport D. C., *The event scheduling problem in molecular dynamics simulation*, J. Comput. Phys., **34**, 184-201, (1980).
- [Rap95] Rapaport D. C., *The art of molecular dynamics simulation*. Cambridge University Press, Gran Bretaña. (1995).

- [Ray72] Ray J. R., *Am. J. Phys.*, **40**, 179, (1972).
- [Ray81] Ray J. R., Graben H. W. y Haile J. M., *Statistical mechanics of the isenthalpic-isobaric ensemble*, *Nuovo Cimento*, **64B**, 191-206, (1981).
- [Rea50] Read W. T. y Shockley W., *Phys. Rev.*, **78**, 275, (1950).
- [Rom97] Romeu D., Beltrán del Río L., Aragón J.L. y Gómez A., en *New horizons in quasicrystals: research and applications*, ed. Goldman A. I., Sordelet D. J., Thiel P. A. y Dubois J. M., World scientific, Singapore. p. 673, (1997).
- [Rom99] Romeu D., Beltrán-del-Río L., Aragón J. L. y Gómez A., *Detailed atomic structure of arbitrary fcc [100] twist grain boundaries*, *Physical Review B*, Vol. **59**. Num. 7 , 5134, (1999).
- [Ros13a] Rosenhain W. y Ewen D., *J. Inst. Metals*. **10**, 125, (1913).
- [Ros13b] Rosenhain W. y Humphrey J. C. W., *J. Iron Steel Inst.*, **87**, 219, (1913).
- [Ros83] Rose J. H., Smith J. R., Ferrante J., *Universal features of bonding in metals*, *Phys. Rev. B*, **28**, vol 4, 1835, (1983).
- [Ros84] Rose J. H., Smith J. R., Guiney F., Ferrante J., *Universal features of the equation of state of metals*, *Phys. Rev. B*, **29**, vol. 6, 2963, (1984).
- [Rut65] Rutter J. W. y Aust K. T., *Acta Met.*, **13**. 181, (1965).
- [San76] Sangster M. J. L. y Dixon M., *Interionic potentials in alkali halides and their use in simulation of molten salts*, *Adv. Phys.* **25**, 247-342, (1976).
- [Sar68] Sargent C.M., *T.M.S.-A.I.M.E.*, **242**, 1188, (1968).
- [Sch70] Schober T. y Balluffi R. W., *Phil. Mag.*, **21**, 109, (1970b).
- [Sch73] Schofield P., *Comp. Phys. Comm*, **5**, 17, (1973).
- [Sch95] Schildt H., *C, The complete reference*, 3a Ed., Osborne McGraw-Hill, E.U.A., (1995),
- [Sim76] Simpson C. J., Winegard W. C. y Aust K. T., *Grain boundary Mobility and Boundary-Solute Interactions*, cap. 6 en [Cha76] (1972).
- [Sin83] Singer K., *A fast method for the evaluation of square roots in computer simulation*, *CCP5 Quarterly*. **8**. 47-8, (1983).
- [Ste80] Steele W. A. y Streett W. B., *Computer simulations of dense molecular fluids. I. Time-dependent statistical properties of single diatomic molecules*, *Mol. Phys.* **39**, 279, (1980).
- [Sti85] Stillinger F. y Weber T. A., *Phys. Rev. B*. **31**, 5262, (1985).
- [Sto73] Stoddard S. D. y Ford J., *Numerical experiments on the stochastic behaviour of a Lennard-Jones gas system*. *Phys. Rev*, **A8**, 1504-12, (1973).
- [Str78a] Streett W. B., Tildesley D. J. y Saville G., *Multiple time step methods and an improved potential function for molecular dynamics simulations of molecular liquids*, en *Computer modelling of matter* (ed. P. Lykos). ACS Symposium Series Vol. 86, pp. 144-58, American Chemical Society. Washington (1978).
- [Str78b] Streett W. B., Tildesley D. J. y Saville G., *Multiple time step methods in molecular dynamics*, *Mol. Phys.*. **35**. 639-48, (1978).
- [Str94] Stroustrup B. y Ellis M. A., *C++. Manual de Referencia con anotaciones*. Addison-Wesley, E.U.A., (1994).
- [Sut87] Sutton A. P. y Balluffi R. W., *Acta Metall.*. **35**. 2177, (1987)
- [Sut93] Sutton A. P., *Electronic structure of materials*. Oxford science, (1993).
- [Swi84] Swindoll R. D. y Haile J. M., *A multiple time step method for calculating the solubility of gases in liquids and the hydrophobic hydration of inert gas atoms in aqueous solution*. *J. Phys. Chem.*, **88**, 6548-56, (1984).

- [Swo82] Swope W. C., Andersen H. C., Berens P. H. y Wilson K. R., *A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters*, J. Chem. Phys., **76**, 637-49, (1982).
- [Ter88] Tersoff J., Phys. Rev. B, **37**, 6991, (1988).
- [Tho04] Thomson W. (Lord Kelvin), *Baltimore lectures on Molecular Dynamics and the Wave Theory of Light*, C. J. Clay and Sons (Cambridge University Press), London, (1904). Reprinted as *Kelvin's Baltimore Lectures and Modern Theoretical Physics*, Kargon R. y Achinstein P., Eds., MIT Press, Cambridge, MA, 1987.
- [Tho83] Thompson S. M., *Use of neighbour lists in molecular dynamics*, CCP5 Quarterly, **8**, 20-8, (1983).
- [Til81] Tildesley D. J. y Madden P. A., *An effective pair potential for liquid carbon disulphide*, Mol. Phys., **42**, 1137-56, (1981).
- [Tox82] Toxvaerd S., *A new algorithm for molecular dynamics calculations*, J. Comput. Phys., **47**, 444-51, (1982).
- [Tur54] Turnbull D. y Hoffman R., *Acta Met.*, **2**, 419, (1954).
- [Ver67] Verlet L., *Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules*. Phys. Rev. **159**, 98-103 (1967).
- [Vic72] Viellard-Baron J.E., *Phase transitions of the classical hard ellipse system*, J. Chem. Phys. **56**, 4729-44, (1972).
- [Vit80] Vitek V. y Sutton A.P., *Scripta Metallurgica*, **14**, 129, (129); Phil. Trans, R. Soc. A, **309**, 37, (1983).
- [Vit96] Vitek V. *Pair Potentials in atomistic simulations*, En MRS bulletin/February 1996. p. 20
- [Vot87] Voter A. F. y Chen S. P., en *Characterization of defects in materials*, editado por Siegal R. W., Wectman J. R. y Sinclair R., MRS Symposia Proceedings No. 82, (Materials Research Society, Pittsburgh, 1987), p. 175.
- [Vot96] Voter A. F., *Interatomic potentials for atomistic simulations*, en MRS Bulletin/February 1996.p. 17
- [War74] Warrington D. H. y Grimmer H., Phil. Mag., **30**, 461, (1974).
- [Wei69] Weins M., Chalmers B., Gleiter H. y Ashby M., *Scripta Met.* **3**, 602. (1969).
- [Wei70] Weins M., Gleiter H. y Chalmers B., *Scripta Met.*, **4**, 235, (1970).
- [Wei72] Weins M. y Weins J. J., Phil. Mag., **26**, 885, (1972).
- [Wol92] Wolf D. y Merkle K. L., en *Materials Interfaces, Atomic Level Structure and Properties*, editado por by D. Wolf y Yip S., Chapman and Hall, Londres, (1992).
- [Woo57] Wood, W. W. y Jacobson J.D., *Preliminary results from a recalculation of the monte carlo equation of state of hard spheres*, J. Chem. Phys. **27** (1957) 1207-1208.
- [Ye96] Ye Y. Y. y Cai J., *Simple analytical embedded-atom-potential model including a long-range force for fcc metals and their alloys*, Phys. Rev. B., **54**, 8398-8410, (1996).
- [Yip] Yip S. en [Cic86] pág 523.