

Seminar

Numerik Gewöhnlicher Differentialgleichungen

Leitung Prof. Peter Bastian
WS 2010/11

Ausarbeitung

Parallele Lösung großer ODE-Systeme
der Moleküldynamik

Dominik Ducati

Inhalt

I Einführung und Grundlagen der Moleküldynamik	S 3
II. Das Lennard-Jones Potential und die Anpassung der Potentialfunktion	S 4
II.1 Das Lennard-Jones Potential	S 4
II.2 Anpassung des Potentials	S 5
II.2.1 Flüssigkeiten A-B	S 5
II.2.2 Elektrische Ladung	S 5
II.2.3 Inermolekulare Bindungen	S 5
III. Berechnung der Kräfte	S 6
IV. Der Linked-Cell Algorithmus	S 7
IV.1 Grundlage	S 8
IV.2 Mögliche Implementierung des Linked-Cell Algorithmus	S 8
IV.3 Pseudo-Code des Linked-Cell Algorithmus	S 9
IV.4 Verbesserungen des Linked-Cell Algorithmus	S 10
V. Parallelisierung	S 10
V.1 Schwierigkeiten und Lösungen bei der Parallelisierung	S 11
V.2 Der Linked-Cell Algorithmus und Parallelisierung	S 12
V.3 Behandlung der Borderzone	S 12
V.4 Geschwindigkeit	S 13
Quellenangabe	S 14

I Einführung und Grundlagen

Die Moleküldynamik (MD) ist eine Form der Computersimulation, bei der Wechselwirkungen zwischen Atomen und Molekülen berechnet werden. Mit Hilfe von physikalischen Approximationen bekommt man einen Einblick in die Bewegung der Partikel. Diese Technik kann man salopp als ein virtuelles Mikroskop verstehen, dessen Auflösung alle tatsächlichen Geräte weit übertrifft. Ein Begriff den man häufig im Zusammenhang mit der Moleküldynamik findet ist „Virtuelles Experiment“, er ist gleichermaßen einprägsam wie treffend. Eine große Rolle spielt die Moleküldynamik bei der Simulation von Flüssigkeiten wie z.B. Wasser, da man dynamische und strukturelle Eigenschaften bei Bedingungen berechnen kann, die experimentell schwer zu realisieren sind, beispielsweise extremer Druck und Temperatur. Bei der Moleküldynamik sind aufgrund der Teilchenanzahl sehr viele Berechnungen und enormer Speicher notwendig. Für komplizierte Simulationen ist Parallelisierung notwendig.

I.1 Grundlagen

Zu Anfang werden die zu untersuchenden Teilchen in das zu betrachtende Volumenelement gefüllt. Für jedes Teilchen werden die Kräfte, die aufgrund seiner Nachbarn auf es wirken berechnet. Anschließend werden die Teilchen entsprechend dieser Kräfte in sehr kleinen Zeitschritten bewegt. In der Molekular Dynamik geht es i. A. um das Lösen der molekularen Bewegungsgleichungen. Die Kräfte \vec{f}_i die auf jedes Teilchen wirken, müssen für jeden Zeitschritt basierend auf Paarpotentialen berechnet werden. Wir betrachten die Teilchen somit paarweise, wobei es zu beachten gilt, dass jedes Paar nur einmal berechnet wird. Letzteres folgt aus dem dritten Newtonschen Gesetz der klassischen Mechanik: actio gleich reactio. Gewöhnlich werden in der Moleküldynamik die Hamilton-Bewegungsgleichungen den klassischen Newtonschen Bewegungsgleichungen vorgezogen. Es hat zwei Vorteile: Zum einen bilden sie ein System von gewöhnlichen Differentialgleichungen erster Ordnung, zum anderen ist es aufgrund der verallgemeinerten Ortskoordinaten \vec{q} einfacher Zwangsbedingungen, wie z.B. eingeschränkte Freiheitsgrade zu berücksichtigen. Die Hamiltonfunktion kann man als Energiefunktion aus der Summe der kinetischen und der potentiellen Energie ansehen.

$$\text{Hamiltonfunktion: } H(\vec{p}, \vec{q}, t) = T(\vec{p}, t) + V(\vec{q}) \quad \text{wobei: } \vec{p} := \text{kan. Impuls, } \vec{q} := \text{Ortskoordinate, } t := \text{Zeit}$$

Mit dem kanonischen Impuls \vec{p} kann man die kinetische Energie wie folgt schreiben:

$$\text{Kinetische Energie: } T(\vec{p}^N) = \sum_{i=1}^N \frac{|\vec{p}_i|^2}{2m_i} \quad \text{wobei: } m := \text{Masse}$$

Die Kräfte \vec{f}_i berechnen sich aus dem Gradienten des Potentials $V(\vec{q})$

$$\text{Es gilt: } \vec{f}_i = -\frac{\partial}{\partial \vec{r}_i} V = -\text{grad}[V(\vec{q})]$$

Damit ergeben sich die Hamilton-Bewegungsgleichungen.

$$\text{Bewegungsgleichungen: } \dot{\vec{q}} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}_i}{m_i} \quad \dot{\vec{p}}_i = -\frac{\partial H}{\partial \vec{q}} = \text{grad}(V) = \vec{f}_i$$

Für die Lösung dieser Gleichungen existieren viele Verfahren. In der Literatur findet man häufig das Verlet Verfahren, das Speed-Verlet Verfahren und das Leap-Frog Verfahren.

II. Das Lennard-Jones Potential und die Anpassung der Potentialfunktion

Ein System ist komplett festgelegt durch seine Hamiltonfunktion $H=H_0+H_1$ wobei H_0 der interne Part der Hamiltonfunktion ist, gegeben durch:

$$H_0 = \sum_{i=1}^N \frac{|\vec{p}_i|^2}{2m_i} + \sum_{i<j}^N V(\vec{r}_i, \vec{r}_j) + \sum_{i<j}^N V^{(3)}(\vec{r}_i, \vec{r}_j, \vec{r}_k) + \dots$$

wobei V das Potential zwischen zwei Teilchen und $V^{(3)}$ das Potential zwischen drei Teilchen usw. angibt. H_1 Ist der externe Part der Hamiltonfunktion, welcher Zeitabhängigkeiten oder äußere Kräfte wie z.B. eine Gefäßwand beschreiben kann.

Gewöhnlich konzentriert man sich in den Experimenten auf die Paarpotentiale V und vernachlässigt sowohl H_1 sowie die Potentiale zwischen 3 und mehr Teilchen.

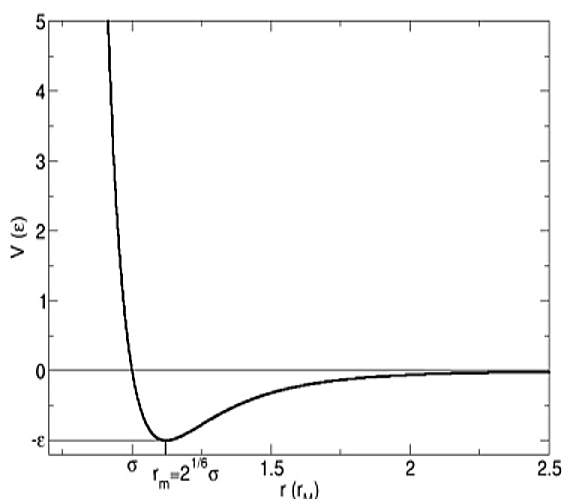
II.1 Das Lennard-Jones Potential

In der Moleküldynamik benutzt man häufig das Lennard-Jones-Potential. Obwohl es ein simples Modell ist, gibt es eine gute Näherung der Wechselwirkung zwischen ungeladenen, nicht chemisch aneinander gebundenen Teilchen. Wir betrachten im weiteren Verlauf nur „short range“ Potentiale, d.h. ab einem gewissen Abstand gehen wir von einer nur sehr geringen Einwirkung der Teilchen aufeinander aus. Das Lennard-Jones Potential ist gegeben durch die Formel:

$$V_{ij}^{LJ}(r_{ij}) = 4\epsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right)$$

wobei $r_{ij} = |r_{ij}| = |r_j - r_i|$ der Abstand zwischen 2 Atomen i und j, σ Atomdurchmesser und ϵ Energieparameter ist. Der Atomdurchmesser sowie der Energieparameter sind atomspezifische Konstanten.

Potentialkurve des L.-J. Potentials



Der Tiefpunkt (Gleichgewichtslage) liegt bei $r = 2\sigma$

Rechts vom Tiefpunkt herrscht Anziehung zwischen den Teilchen. Man erkennt deutlich das das Potential sehr schnell gegen 0 geht mit zunehmendem Abstand.

Links vom Tiefpunkt herrscht Abstoßung zwischen den Teilchen.

Bild: Quelle: Wikipedia

II.2 Anpassung des Potentials

Eine schöne Eigenschaft des physikalischen Potentials ist, dass man es erweitern bzw. an das Problem anpassen kann. Sind die Atome beispielsweise geladen, addiert man das Coulomb Potential mit den entsprechenden Ladungen. Bei nicht punktförmigen Teilchen ist es möglich, den Term um das Rotationspotential zu erweitern. Das Lennard-Jones Potential hat darüber hinaus die Eigenschaft, dass man es durch einfache Kombinationsregeln der Konstanten ϵ und σ auf verschiedenartige Teilchen anwenden kann.

Es ist sogar möglich mit Hilfe der „short-range“ Potentiale ganze Moleküle aus Atomen aufzubauen. Bei diesem Prozess kommen eine ganze Reihe weiterer Potentiale zum Zug. Man muss hierbei zusätzlich chemische Bindungen berücksichtigen und für die Berechnung der innermolekularen Winkel reicht es nicht mehr die Atome nur paarweise zu betrachten.

II.2.1 Flüssigkeiten A-B

Betrachtet man eine Mischung bestehend aus zwei Komponenten A und B so können 3 verschiedene Paare entstehen. Bei der Paarbildung A-A und B-B treten keinerlei Probleme mit dem Lennard-Jones Potential auf. Bei der Paarbildung A-B allerdings, müssen der Atomdurchmesser σ und der Energieparameter ϵ angepasst werden.

$$\text{Lorentz-Berthelot Kombinations Regeln : } \sigma_{AB} = \eta_{AB} \frac{\sigma_A + \sigma_B}{2}, \quad \epsilon_{AB} = \xi_{AB} \sqrt{\epsilon_A \epsilon_B}$$

wobei: η meist auf 1 gesetzt und der Faktor ξ nahe der 1 liegt

II.2.2 Elektrische Ladung

Falls elektrische Ladung in der Simulation berücksichtigt werden soll, müssen wir das zugehörige Coulomb-Potential addieren.

$$V_c(r) = \frac{1}{4\pi\epsilon_0} \frac{Q_i Q_j}{|\vec{r}_i - \vec{r}_j|}$$

wobei : Q_i die Ladung des Teilchens i, Q_j die Ladung des Teilchens j
 ϵ_0 Elektrische Feldkonstante im Vakuum (Permittivität des Vakuums)
 $|\vec{r}_i - \vec{r}_j|$ der Abstand zwischen den Teilchen

II.2.3 Innermolekulare Bindungen

Moleküle können mit Hilfe von „short-range“ Potentialen aufgebaut werden. Hierbei werden noch eine ganze Reihe weiterer Potentiale benötigt. Das Bild auf der nächsten Seite illustriert sehr schön einige der meist benutzten Potentiale.

Quelle: *Classical Molecular Dynamics - Godehard Sutmann*

6-9 van der Waals		$u_{ij}(r_{ij}) = 4\epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^9 - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right)$
6-12 van der Waals		$u_{ij}(r_{ij}) = 4\epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right)$
Electrostatic		$u_{ij}(r_{ij}) = \frac{q_i q_j}{r_{ij}}$
Quadratic bond-stretching		$u_{ij}^s(r_{ij}) = \frac{1}{2} k_{ij} (r_{ij} - b_{ij})^2$
Morse bond-stretching		$u_{ij}(r_{ij}) = k(1 - e^{-a(r_{ij} - r_0)})^2$
Bond-bending		$u_{ij}^b(\vartheta_{ijk}) = \frac{1}{2} k_{ijk} (\vartheta_{ijk} - \vartheta_{ijk}^0)^2$
Improper dihedrals		$u_{ij}^{id}(\xi_{ijkl}) = \frac{1}{2} k_{ijkl} (\xi_{ijkl} - \xi_0)^2$
Proper dihedrals		$u_{ij}^{pd}(\varphi_{ijkl}) = k_\varphi (1 + \cos(n\varphi_{ijkl} - \varphi_0))$

Figure 2. Typical examples for potential terms as used in common force-fields.

III. Berechnung der Kräfte

Hat man nun die Potentiale bestimmt ist der nächste Schritt die Berechnung der Kräfte. Diese Rechnung ist der kostspieligste Part der ganzen Simulation. Es hat sich gezeigt das etwa 90% der Rechenzeit der gesamten Simulation dafür in Anspruch genommen wird. Wir berechnen die Kräfte mit Hilfe des Gradienten aus den Potentialen. Ein große Hilfe ist, dass die Potentiale in der Regel analytisch ableitbar sind.

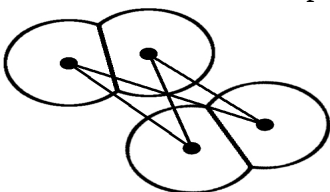
Es wird i. A. immer mit der analytischen Ableitung gearbeitet wenn es möglich ist.

Die Lösung ist exakter und man ist schneller, da der Rechenaufwand nicht so groß ist.

Beispiel:

Quelle: *Aspects of a Parallel Molecular Dynamics Software for Nano-Fluidics*
- Martin Bernreuther¹, Martin Buchholz², and Hans-Joachim Bungartz²

Als Beispiel dient hier das Quadrupol „Two Center Lennard-Jones (2CLJQ)“ - Molekül Modell. Man kann mit diesem Modell Substanzen wie z.B. Sauerstoff (O_2) oder Stickstoff (N_2) darstellen. Es wird uns später bei der Beurtelung der Geschwindigkeiten noch einmal begegnen.



Es wird hier nur die Beziehung zwischen 2 Molekülen angegeben. Genauer gesagt die Kräfte zwischen den beiden Zentren der Moleküle.

Potentialfunktion: Man sieht hier das Lennard-Jones Potential erweitert um das Coulomb-Potential und eine Rotationsfunktion.

$$V_{ij}^{2CLJQ} = 4 \sum_{k=1}^2 \sum_{l=1}^2 \epsilon_{kl} \left(\left(\frac{\sigma_{kl}^2}{r_{kl}^2} \right)^6 - \left(\frac{\sigma_{kl}^2}{r_{kl}^2} \right)^3 \right) \frac{\vec{r}_{kl}}{r_{kl}^2} + \frac{3}{4} \frac{Q_i Q_j}{r_{ij}^6} f(\omega_i, \omega_j)$$

Kraftgleichung: Analytische Ableitung aus der Potentialfunktion.

$$\vec{F}_{ij}^{2CLJQ} = 24 \sum_{k=1}^2 \sum_{l=1}^2 \epsilon_{kl} \left(2 \left(\frac{\sigma_{kl}^2}{r_{kl}^2} \right)^6 - \left(\frac{\sigma_{kl}^2}{r_{kl}^2} \right)^3 \right) \frac{\vec{r}_{kl}}{r_{kl}^2} + \frac{15}{4} \frac{Q_i Q_j}{r_{ij}^6} f(\omega_i, \omega_j) \frac{\vec{r}_{ij}}{|r_{ij}|}$$

IV Der Linked-Cell Algorithmus

IV.1 Grundlage

Geht man von einem Versuch mit N Teilchen aus, so ergibt sich die Anzahl der Paare wie folgt :

$$\binom{N}{2} = \frac{1}{2} N(N-1) \quad \text{Möglichkeiten der Paarbildung}$$

Dies entspricht einem Aufwand von $O(N^2)$. Wir können benutzen das die Kräfte die die Teilchen aufeinander wirken bei dem Lennard-Jones Potential sehr schnell verschwindend gering werden. Man kann dies anhand des Graphen sehr schön erkennen. Rechts der Gleichgewichtslage (Tiefpunkt des Potentials) geht der Graph der Kurve schnell gegen 0 mit wachsendem Abstand. Man setzt aus diesem Grund alle Kräfte gleich Null für einen Abstand $r > r_c$. Die konstante r_c wir hierbei Cutoff-Distance (bzw. Cutoff-Radius) genannt.

Der linked-Cell Algorithmus ist zum schnellen Auffinden aller Atompaaire innerhalb einer Cutoff Distanz entwickelt worden. Die Idee des Algorithmus ist den Simulationsraum in kubische Zellen einzuteilen und die Teilchen der jeweiligen Zelle zuzuordnen. Im 3-dimensionalen Raum hat demnach eine Zelle 26 Nachbarzellen. Man kann nun alle Partner eines Teilchens finden in dem man es erst mit allen Teilchen der eigenen Zelle und anschließend mit denen aller Nachbarzellen vergleicht.

Arbeitsweise:

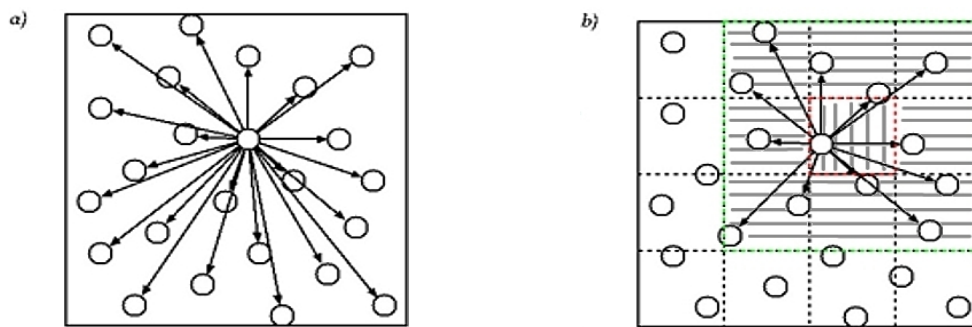
- 1) Teile den Simulationsraum in kubische Würfel (Zellen) mit Seitenlänge r_c
- 2) Ordne die Teilchen der jeweiligen Zelle zu

Finden von Atompaairen:

- 1) Berechne Kräfte zwischen dem i-ten Atom und allen in der selben Zelle befindlichen Atomen
- 2) Berechne Kräfte zwischen dem i-ten Atom und allen Atomen in der Nachbarzelle

Den Unterschied zwischen einer Naiven Vorgehensweise und dem Linked Cell-Algorithmus verdeutlicht die nachfolgende Abbildung.

Originales Bild: Quelle: Wikipedia



- a) Naive Vorgehensweise - Die Berechnung der paarweisen Interaktion eines Teilchens mit allen anderen
- b) Linked Cell - Unterteilung des Raumes mit „Cutoff - radius“ r_c und anschließende Berechnung der Interaktion eines Teilchens mit allen Teilchen der selben Zelle (senkrecht) und allen benachbarten Zellen (waagrecht)

Der Linked-Cell Algorithmus arbeitet mit dem Aufwand $O(N)$. Man kann dies leicht nachweisen mit einer Abschätzung nach oben.

Angenommen die Gesamtanzahl unsere Atome sei N , weiterhin liegt eine homogene Verteilung mit n Atomen in einer Zelle vor. Ein Atom hat demnach $(n-1)$ Vergleiche in der eigenen Zelle und zusätzlich weitere $26n$ Vergleiche mit den Atomen der Nachbarzellen.

Damit ist die Anzahl der Vergleiche für eine Zelle:

$$n[(n-1)+26n]=n(27n-1)$$

Weiter gilt für die Anzahl der Vergleiche für alle Zellen:

$$(N/n)*n(27n-1)=(27n-1)N$$

Wobei N/n der Gesamtanzahl der Zellen entspricht. Da n sehr klein im Verhältnis zu N ist folgt das der Aufwand $O(N)$ beträgt.

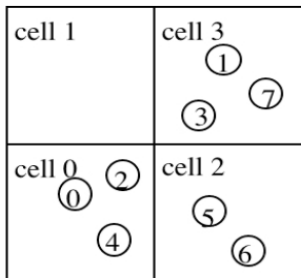
Nach physikalischem Standpunkt begeht man mit dem Abschneiden des Potentials ab einem gewissen Abstand einen Fehler. Dieser Ungenauigkeit kann man Rechnung tragen indem man an die Potentialfunktion bzw. der Kräftegleichung einen Korrekturterm anhängt.

IV.2 Mögliche Implementierung des Linked-Cell Algorithmus

Eine effiziente Möglichkeit der Implementation des Linked-Cell Algorithmus funktioniert mit Hilfe zweier Arrays. Das eine Array, der Header enthält genau ein Atom aus jeder Zelle. Man kann ihn sich als eine Art Hash-Table vorstellen, bei der ein Vertreter auf alle anderen Vertreter der Zelle verweist. Das Zweite Array enthält demnach alle anderen Atome, die nicht im Header enthalten sind.

Quelle: cacs.usc.edu/education/cs596/01-1LinkedListCell.pdf

- Arbeitsweise: - Die Atome werden durchnummeriert
 - im Header *head* wird immer das Atom mit der höchsten Nummer in der Zelle gespeichert (oder Empty (-1) falls keines vorhanden ist).
 - in dem Array *lscl* werden alle Atome in bestimmter Reihenfolge gespeichert



	0	1	2	3				
head	4	E	6	7				
lscl	E	E	0	1	2	E	5	3

N : Anzahl der Atome
 NC : Anzahl der Zellen

lscl[N] : Eine Array implementation aller „Linked - Lists“

head[NC] : head[c] enthält den Index des ersten Atoms in der c-ten Zelle

lscl[i] enthält den Atomindex auf den das i-te Atom zeigt

head[c] = Empty (= -1) wenn es kein Atom in der Zelle gibt



Beispiel:

- In Zelle 3 (d.h. Position 3 im header) ist Atom Nr. 7 gespeichert
- an der Stelle 7 im Array **lscl** ist das Atom Nr. 3 gespeichert
- an der Stelle 3 im Array **lscl** ist das Atom Nr. 1 gespeichert
- an der Stelle 1 im Array **lscl** ist schließlich Empty gespeichert

Wir erhalten also in Zelle 3: 7 → 3 → 1

IV.3 Pseudo-Code des Linked-Cell Algorithmus

```

for ( c = 0; c < Lx * Ly * Lz; c++)           // Zellenanzahl in x,y,z Richtung
    head[c] = Empty;                          // Zurücksetzen des Headers
for ( i = 0; i < AnzahlAtome; i++) {         // Gehe durch alle Atome

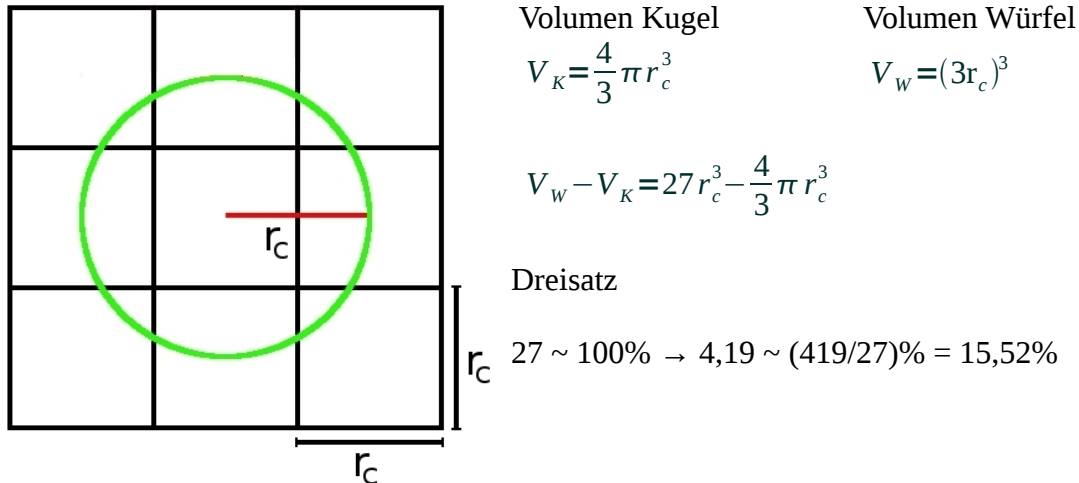
    mc_x = r_x[i] / rc_x;                     // Hilfsvariable = x-Koordinate[Atom] / rc
    mc_y = r_y[i] / rc_y;                     //      ...
    mc_z = r_z[i] / rz_z;                     //      ...

    c = mc_x * Ly * Lz                        // Umwandeln von mc in einen skalaren
        + mc_y * Lz + mc_z                    // Zellenindex

    lscl[i] = head[c];                        // Index des letzten Vorgängers (oder Empty)
    head[c] = i;                              // Der letzte geht in der header
}
  
```

IV.4 Verbesserungen des Linked-Cell Algorithmus

Die Zellengröße r_c zu wählen ist nicht sehr gut geeignet. Deutlich wird dies wenn man sich ein Atom als Mittelpunkt wählt und nun das Volumen der Kugel mit Radius r_c mit dem Volumen der betreffenden Zellen vergleicht. In der Skizze sieht man einen 2-dimensionalen Querschnitt der Szene.



Die Kugel um das Atom überdeckt nur etwa 16% des Volumens der Zellen. Da man die Kräfte gleich Null für einen Abstand $r > r_c$ setzt, ist es möglich durch eine geschickte Wahl der Zellengröße viel Rechenleistung ein zu sparen. Man ist dazu übergegangen eine Zellengröße von r_c/t mit $t \in \mathbb{R}^{+6}$ zu implementieren. Der Vorteil ist, dass man dadurch eine höhere Flexibilität und ebenso Auflösung bekommt (Für $t \rightarrow \infty$ hätte man eine ideale Kugel). Das Problem ist, wenn die Zellen zu klein werden, müssen sehr viele Leere Zellen getestet werden, was wiederum die Laufzeit erheblich erhöht.

Für Flüssigkeiten hat sich in der Praxis ein Wert von $t = 2$ als sehr gut gezeigt.

V Parallelisierung

Allgemein versteht man unter der Parallelisierung, das Aufteilen des Problems auf mehrere Computer. Gewöhnlich arbeitet man mit Clustern, einer Anzahl von vernetzten Computern die von außen als ein Computer angesehen werden können. Durch die Verwendung eines Clusters erhält man bei der Moleküldynamik einen enormen Geschwindigkeitszuwachs bei den Simulationen aufgrund der erhöhten Rechenleistung. Ein weiterer Effekt den der Einsatz mehrerer Computer mit sich bringt ist der Zuwachs an Arbeitsspeicher, auf den im Vergleich zu einer Festplatte schnell zugegriffen werden kann.

Die Arbeitsweise des Linked-Cell Algorithmus eignet sich hervorragend für die Parallelisierung, da man den Experimentierraum ohnehin in Kubische Würfel unterteilt hat. Es bietet sich geradezu an das Volumen auf die verschiedenen Prozessoren zu verteilen.

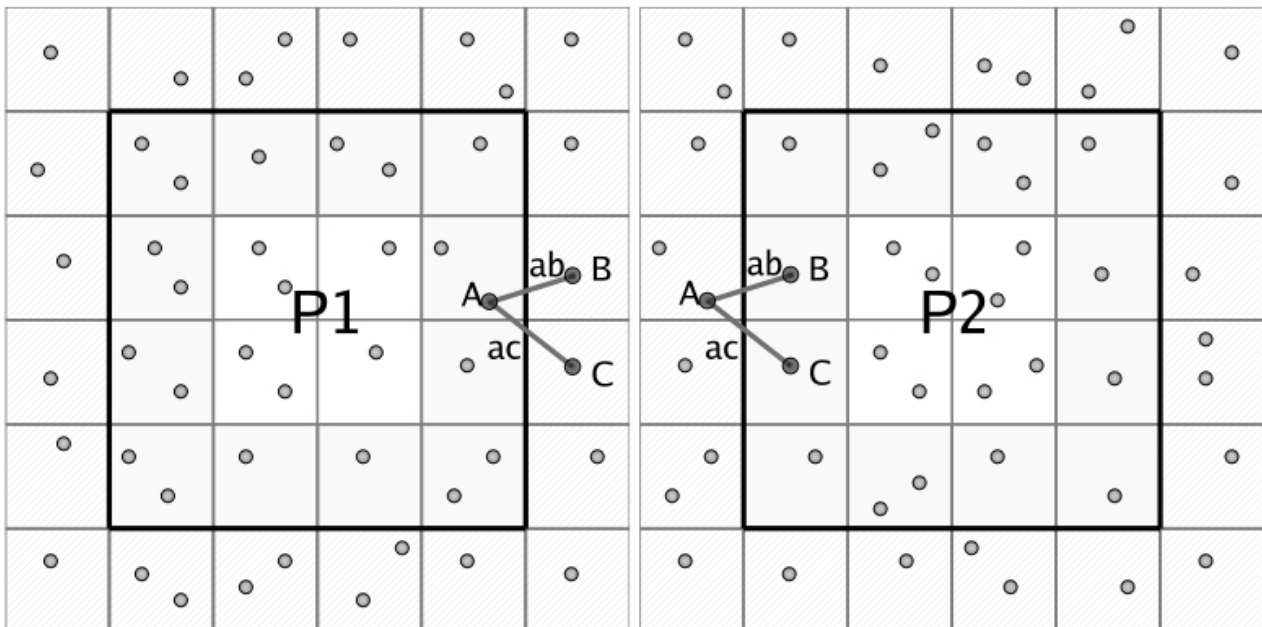
V.1 Schwierigkeiten und Lösungen bei der Parallelisierung

Bei der Parallelisierung gilt es elegante Lösungen für die auftretenden Schwierigkeiten zu finden. Eines der Probleme ist, dass sich die Randatome zwischen 2 benachbarten, in verschiedene Computer aufgeteilten Volumina gegenseitig beeinflussen. Eine besondere Beachtung muss man ebenso dem Rand unseres Experimentierraums, also dem Gefäßrand, widmen. In den meisten Simulationen ist man daran interessiert ein unendliches Gefäß zu modellieren, also nicht am Rand des Experimentierraums beschränkt zu sein. Eine weitere Hürde die bei nahezu allen Problemen, die man mit Parallelisierung angehen möchte, auftritt ist die Kommunikation zwischen den Rechnern.

Sie darf nicht ausarten, da es beispielsweise unzweckmäßig ist, wenn ein großer Teil der CPUs ständig auf Ergebnisse eines anderen Rechners warten muss. Man möchte schließlich einen möglichst hohen Zeitgewinn aus der Parallelisierung erhalten.

Der gegenseitigen Beeinflussung der Volumina trägt man Rechnung, indem jeder Computer zusätzlich zu dem ihm zugeteilten Volumen des Experimentierraums eine sogenannte Border Region (auch Border Zone genannt), erhält. In dieser Border Region befinden sich Kopien des Randes der angrenzenden Volumina. Bedingt durch die Cutoff Distance hat diese Border Zone die Größe r_c , somit ist sichergestellt das auch alle zur Berechnung relevanten Atome berücksichtigt werden. Diese Methode eignet sich ebenso hervorragend für die Modellierung eines unendlichen Simulationsraums. Um dies zu erreichen werden die Border Regions des Randes des Experimentierraumes mit den Atomen des Randes der gegenüberliegenden Seite aufgefüllt.

Quelle: *Aspects of a Parallel Molecular Dynamics Software for Nano-Fluidics*
- Martin Bernreuther, Martin Buchholz, Hans-Joachim Bungartz



Dieses Schaubild verdeutlicht die Volumenzugehörigkeit zu P1 und P2 sowie die Border Region. Auf der nächsten Seite wird die Behandlung der Teilchen A,B,C in der Border Region geklärt.

V.2 Der Linked-Cell Algorithmus und Parallelisierung

Die Frage, die sich grundsätzlich bei Parallelisierung stellt ist, ob sich der Algorithmus überhaupt parallelisieren lässt. Es liegt nahe, wenn man mehrere Computer zur Verfügung hat, die Teilchen auf die Computer zu verteilen. Wir beschäftigen uns also mit einer durch den L-C Algorithmus bedingten räumlichen Zerlegung .

Der Experimentraum wird unterteilt, so dass jeder Computer ein kubisches Volumen erhält. Jeder Computer berechnet somit die Kräfte und neue Positionen ausschließlich für einen Teil der Atome. Ein weiterer Gewinn ist, dass jeder Computer nur seinen Teil der Atome speichern muss. Da hier ein Cluster betrachtet wird, kann man also davon ausgehen, dass mehrere voneinander unabhängige Arbeitsspeicher zur Verfügung stehen. Das bedeutet insbesondere bei großen Datenmengen einen enormen Vorteil aufgrund der schnellen Speicherzugriffe. Es ist noch die Frage der Kommunikation zu klären. Nach jedem Zeitschritt müssen die Daten synchronisiert werden. Die Border Regions sind bedingt durch unsere Cutoff-Distance relativ klein und enthalten dadurch nicht so viele Atome, was den Datentransfer und Speicheraufwand handhabbar macht. Jeder Rechner muss nur einen Teil des Experimentraumes speichern und nur die Border-Zone muss mit anderen Rechnern synchronisiert werden. Endgültig beantworten kann man die Frage nach der Parallelisierbarkeit und dem dadurch erhaltenen Gewinn durch Geschwindigkeitsmessungen und Vergleiche.

V.3 Behandlung der Border Zone

Für die Berechnung der Border Zone gibt es verschiedene Herangehensweisen. Es gibt die Möglichkeit, dass nur einer der beiden Computer die Border Region berechnet und seine Ergebnisse dem anderen Rechner zur Verfügung stellt. Da man davon ausgehen kann, dass die Border Zone sehr klein ist und damit wenige Atome enthält, kann man sie auch durch beide Computer berechnen lassen. In dem Paper „*Aspects of a Parallel Molecular Dynamics Software for Nano-Fluidics*“ hat man sich dafür entschieden, dass beide Computer die Border Zone berechnen.

Es wird hier anhand einer „globalen Atomordnung“ (Index) entschieden, von welchem Rechner das Ergebnis genommen wird. Einem Prozess, dem das erste der beiden Teilchen innewohnt gehört das Paar.

Beispiel aus dem Schaubild

Die Zellen wurden von links nach rechts durchnummeriert (Spalte). Jede Zeile beginnt mit dem Index +1 der letzten Zelle der Zeile darunter.

...											
13	14	15	16	...							
1	2	3	4	5	6	7	8	9	10	11	12

C - hat den niedrigsten Index

A - hat den „mittleren“ Index

B - hat den höchsten Index

→ AB gehört dem Computer P1 , AC gehört dem Computer P2

V.4 Geschwindigkeit

Die Geschwindigkeitsmessungen sind dem Paper „*Aspects of a Parallel Molecular Dynamics Software for Nano-Fluidics*“ entnommen. Es wurden Simulationen auf einem 16 Knoten Linux Cluster laufen gelassen. Jeder Knoten hat 8GB RAM und 4 Opteron Prozessoren mit 2.4 Ghz. Es wurde das Quadrupol „*Two Center Lennard-Jones (2CLJQ)*“ - Molekül Modell benutzt. (Potentialfunktion und Kraftgleichung s.o.)

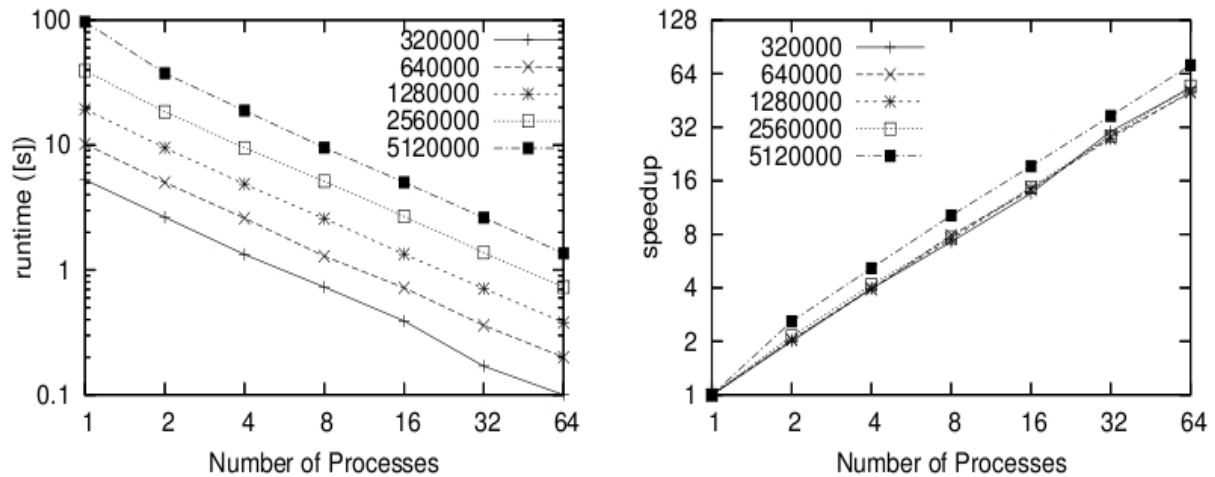


Figure 5. Runtime and speedup for different numbers of particles on up to 64 processors

Auf diesem Bild sind die Zeiten und der Speedup für einen Zeitschritt t angegeben.

Auf dem linken Bild sieht man deutlich das die benötigte Zeit auf einem Einzelnen Prozessor mehr als doppelt so hoch für einen Zeitschritt t ist, wenn die Anzahl der Teilchen von 2.56 Millionen auf 5.12 Million erhöht wird. Das liegt daran das ein einzelner Computer nicht genug Arbeitsspeicher hat. Mit mehreren Prozessoren ist dieses Problem nicht mehr gegeben.

Auf dem rechten Bild sieht man den Geschwindigkeitszuwachs pro Cpu. Betrachtet man genauer die Linie mit den 5.120.000 Atomen so sieht man :

- bei zwei CPU's man schon etwas weniger als die Hälfte der Zeit
- bei 16 CPU's etwa nur 1/16 der Zeit
- bei 64 CPU's nur 1/64

An dem schön linearen Anstieg sieht man, dass das Problem sehr gut parallelisierbar ist. Die benötigte Zeit nimmt mit jedem zugeschalteten Prozessor linear ab. Dies lässt wiederum Rückschlüsse auf die Kommunikation zu. Es entstehen kaum Wartezeiten durch den Datenaustausch im Netzwerk.

Quellenangabe:

Paper

- * Molecular Simulation of Fluids with Short Range Potentials
- Martin Bernreuther¹ and Jadran Vrabec²
- * Aspects of a Parallel Molecular Dynamics Software for Nano-Fluidics
- Martin Bernreuther, Martin Buchholz, Hans-Joachim Bungartz
- * Introduction to Molecular Dynamics Simulation - Michael P. Allen
- * Classical Molecular Dynamics - Godehard Sutmann

Weblink

- * cacs.usc.edu/education/cs596/01-1/LinkedListCell.pdf