

Strukturelle Modellierung
(Masterstudiengang Bioinformatik)

Moleküldynamiksimulation: Theorie

Sommersemester 2012

Peter Güntert

Literatur

- Andrew R. Leach: *Molecular Modelling, Principles and Applications*, Prentice Hall, 2001.
- M. P. Allen & D. J. Tildesley: *Computer Simulation of Liquids*, Clarendon Press, 1987.
- Tamar Schlick: *Molecular Modeling and Simulation*, Springer, 2006.

Computer-Simulation von Molekülen

Strukturelle Modellierung
(Masterstudiengang Bioinformatik)

Computer-Simulation von Molekülen

Sommersemester 2012

Peter Güntert

Computer-Simulation von Molekülen

- Modellierung: QM, klassisch, schematisch
- Klassische Mechanik: Newton, Lagrange, Hamilton Bewegungsgleichungen
- Kartesische Koordinaten, interne Koordinaten
- Statistische Mechanik
- Simulationsmethoden: Systematische Suche, Monte Carlo, MD, Stochastische Dynamik (Langevin), Energieminimierung, Normalmodenanalyse
- Annahmen, Näherungen und Grenzen: Klassische Mechanik, Zeitskala, Systemgröße, Kraftfeld
- Geschichte der MD Simulation

Vielteilchenproblem

	Kristalline Festkörper	Makromoleküle Flüssigkeit	Gasphase
Quantenmechanik ($\sim N^4$)	möglich	(noch?) nicht möglich	möglich
Klassische Mechanik ($\sim N \log N$)	einfach	essentielle Vielteilchensysteme	trivial

←
Reduktion auf wenige
Freiheitsgrade durch
Symmetrie

↑
**Computer
Simulation**

→
Reduktion auf wenige
Teilchen durch
Verdünnung

**Von mikroskopischen Details zu makroskopischen Größen:
Statistische Mechanik**

- Zustandssumme: $Z = \frac{1}{h^{3N} N!} \iint e^{-H(\vec{p}, \vec{q})/kT} d\vec{p}d\vec{q}$
- Hamilton-Funktion/Operator:
 $H(\mathbf{p}, \mathbf{q}) =$ kinetische Energie + potentielle Energie (Wechselwirkung)
- Erzeugung eines Ensembles von Konfigurationen durch Computersimulation
- Mittelung über Ensemble und Zeit liefert makroskopische Größen

Erhaltungsgrößen

- Jeder kontinuierlichen Symmetrie eines mechanischen Systems entspricht eine Erhaltungsgröße. (Satz von Emmy Noether)

Invarianz bezüglich	Erhaltungsgröße
Zeitverschiebung	Energie
Translation	Impuls
Rotation	Drehimpuls

Statistische Mechanik

- **Grundprinzip:** Für ein thermodynamisches System im Gleichgewicht ist die Wahrscheinlichkeit, dass es einen Zustand der Energie E_i annimmt:

$$\text{Prob}(E_i) = \frac{1}{Z} e^{-E_i/kT}, \quad Z = \sum_{i=1}^N e^{-E_i/kT}$$

$k_B = 1.38065 \cdot 10^{-23}$ J/K: Boltzmann-Konstante
 T : absolute Temperatur
 Z : Zustandssumme

- Erwartungswert einer Observablen A : $\langle A \rangle = \frac{1}{Z} \sum_{i=1}^N A_i e^{-E_i/kT}$
- Thermodynamik: freie Energie $F = -k_B T \log Z$

Proteinsimulation: Stufen der Vereinfachung

Modell	Freiheitsgrade	Eliminierte Freiheitsgrade	Untersuchbare Phänomene (Beispiele)
Quantenmechanik	Kerne, Elektronen	Nukleonen	Chemische Reaktionen
Alle Atome, polarisierbar	Atompositionen, Polarisation	Elektronen	Bindung geladener Liganden
Alle Atome	Atompositionen (Protein + Wasser)	Polarisation	Hydratisierung
Alle Proteinatome	Positionen der Proteinatome	Lösungsmittel (Wasser)	Konformation in Gasphase
Gittermodelle	Aminosäuren	Atome	Faltung?

Simulation von Vielteilchensystemen

System: N Atome: $i = 1, \dots, N$, Koordinaten r_i , Massen m_i , Wechselwirkung ("Kraftfeld") $V(r_1, \dots, r_N)$

Methoden:

- Systematische Suche → Ensemble
- Monte Carlo (MC) → Ensemble
- Moleküldynamik (MD) → Trajektorie
- Stochastische Dynamik (SD; Langevin) → Trajektorie
- Energieminimierung (EM) → 1 Konfiguration
- Normalmodenanalyse → 1 Konfiguration + harmonische Bewegung

Vergleich Simulation - Experiment

- **Atomare Eigenschaften**
 - Struktur: Positionen, Distanzen, H-Brücken
 - Mobilität: B-Faktoren, Populationen
 - Dynamik: Vibrationen, Relaxationsraten, Diffusion, Pfade
- **Makroskopische Eigenschaften**
 - Thermodynamische Größen: Druck, Dichte, freie Energie, ...
 - Viskosität, Kompressibilität, Dielektrizitätskonstante

Annahmen, Näherungen und Grenzen

- **Klassische Mechanik:**
 - keine tiefen Temperaturen ($T \approx 300\text{ K}$)
 - Bewegung der Wasserstoffatome
 - keine chemischen Reaktionen
- **Zeitskala von Prozessen:** (im atomaren Bereich)
 - bis ca. $1\ \mu\text{s}$
 - aktivierte Prozesse: möglich
 - essentiell langsame Prozesse: nicht möglich (z.B. Proteinfaltung)
- **Systemgröße:**
 - $N \leq 10^6$ Teilchen
 - keine essentiell makroskopischen Vorgänge (z. B. kritische Phänomene, Phasenübergänge)
- **Kraftfeld:**
 - Ist die atomare Wechselwirkungsfunktion genügend genau zur Vorhersage der gewünschten Eigenschaft?

Strukturelle Modellierung
(Masterstudiengang Bioinformatik)

Moleküldynamiksimulation: Theorie

Sommersemester 2012

Peter Güntert

Literatur

- Andrew R. Leach: *Molecular Modelling, Principles and Applications*, Prentice Hall, 2001.
- M. P. Allen & D. J. Tildesley: *Computer Simulation of Liquids*, Clarendon Press, 1987.
- Tamar Schlick: *Molecular Modeling and Simulation*, Springer, 2006.

MD Simulation: Geschichte

1957: Harte Scheiben (2D)	-
1964: Monoatomare Flüssigkeit	10^{-11} s
1971: Molekulare Flüssigkeit	$5 \times 10^{-12}\text{ s}$
1971: Flüssiges Salz	10^{-11} s
1975: Einfaches kleines Polymer	10^{-11} s
1977: Protein im Vakuum	$2 \times 10^{-11}\text{ s}$
1982: Einfache Membran	$2 \times 10^{-10}\text{ s}$
1983: Protein in Lösung	$2 \times 10^{-11}\text{ s}$
1986: DNA in Lösung	10^{-10} s
1989: Protein-DNA Komplex in Lösung	10^{-10} s
1991: Protein-Protein Komplex in Lösung	10^{-9} s
1998: $1\ \mu\text{s}$ Simulation eines Proteins in Lösung	10^{-6} s

MD Simulation: Personen und Programme



Martin Karplus
(CHARMM)



Peter Kollman (†)
(AMBER)



Herman Berendsen
(GROMOS/GROMACS)

Kraftfelder

Strukturelle Modellierung (Masterstudiengang Bioinformatik)

Kraftfelder

Sommersemester 2012

Peter Güntert

Molekülmodell und Kraftfelder

- Kraftfelder: CHARMM, AMBER, GROMACS, ...
- Energiefunktion: kovalente Bindungen, Bindungswinkel, Diederwinkel, nichtbindende Wechselwirkung (van der Waals und elektrostatisch), H-Brücken
- Bestimmung von Kraftkonstanten
- Paarliten und Cutoffs für nichtbindende Wechselwirkungen

Classical force fields

- **AMBER** (Assisted Model Building and Energy Refinement) - widely used for proteins and DNA
- **CHARMM** (Chemistry at HARvard Molecular Mechanics) - originally developed at Harvard, widely used for both small molecules and macromolecules
- **GROMACS** - The force field optimized for the package of the same name
- **GROMOS** - A force field that comes as part of the GROMOS (GROningen MOlecular Simulation) package, a general-purpose MD package for the study of biomolecular systems
- **OPLS** (Optimized Potential for Liquid Simulations) developed by William L. Jorgensen at Yale University
- **ECEPP/2** - First force field for polypeptide molecules - developed by Harold Scheraga and colleagues

AMBER Force Field

$$E_{\text{pair}} = \sum_{\text{bonds}} K_r (r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_{\text{eq}})^2$$

$$+ \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$$

$$+ \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

Potenzial für kovalente Bindungen

$$\sum_{\text{bonds}} K_r (r - r_{\text{eq}})^2$$

- Summe über alle kovalenten Bindungen im Molekül
- Harmonisches Potential ("Bindung = Feder")
- Bindungen können nicht gebrochen werden
→ keine chemischen Reaktionen
- Parameter, von den Typen der kovalent gebundenen Atome abhängig:
 - K_r Kraftkonstante
 - r_{eq} Gleichgewichtsbindungslänge
- Rechenaufwand proportional zur Molekülgröße
- Bei Rechnung im Torsionswinkelraum nicht nötig

Potenzial für Bindungswinkel

$$\sum_{\text{angles}} K_\theta (\theta - \theta_{\text{eq}})^2$$

- Summe über alle Bindungswinkel im Molekül
- Harmonisches Potential
- Parameter, vom Typ der involvierten Atome abhängig:
 - K_θ Kraftkonstante
 - θ_{eq} Gleichgewichtsbindungs Winkel
- Rechenaufwand proportional zur Molekülgröße
- Bei Rechnung im Torsionswinkelraum nicht nötig

Potenzial für Diederwinkel

$$\sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$$

- Summe über alle Diederwinkel im Molekül
- Periodisches Potential
- Parameter, vom Typ der involvierten Atome abhängig:
 - V_n Kraftkonstante
 - n Anzahl der Energiemaxima
 - γ Position des ersten Energiemaximums
- Rechenaufwand proportional zur Molekülgröße

Potenzial für uneigentliche ("improper") Diederwinkel

$$\sum_{\text{improvers}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$$

- Halten planare Gruppen in einer Ebene (Peptidgruppe, aromatische Ringe usw.)
- Für ausgewählte Quadrupel von Atomen
- Parameter, vom Typ der involvierten Atome abhängig:
 - V_n Kraftkonstante
 - n Anzahl der Energiemaxima
 - γ Position des ersten Energiemaximums
- Rechenaufwand proportional zur Molekülgröße

AMBER Atom Types

Table 1. List of Atom Types*

atom	type	description
carbon	CT	any sp ³ carbon
	C	any carbonyl sp ² carbon
	CA	any aromatic sp ² carbon and (Cε of Arg)
	CM	any sp ² carbon, double bonded
	CC	sp ² aromatic in 5-membered ring with one substituent + next to nitrogen (Cγ in His)
	CV	sp ² aromatic in 5-membered ring next to carbon and lone pair nitrogen (e.g. Cδ in His (δ))
	CW	sp ² aromatic in 5-membered ring next to carbon and NH (e.g. Cδ in His (ε) and in Trp)
	CR	sp ² aromatic in 5-membered ring next to two nitrogens (Cγ and Cε in His)
	CB	sp ² aromatic in junction of 5- and 6-membered rings (Cδ in Trp) and both junction atoms in Ade and Gua
	C*	sp ² aromatic in 5-membered ring next to two carbons (e.g. Cγ in Trp)
	CN	sp ² junction between 5- and 6-membered rings and bonded to CH and NH (Cε in Trp)
	CK	sp ² carbon in 5-membered aromatic between N and N-R (C8 in purines)
	CQ	sp ² carbon in 6-membered ring between lone pair nitrogens (e.g. C2 in purines)

nitrogen	N	sp ³ nitrogen in amides
	NA	sp ³ nitrogen in aromatic rings with hydrogen attached (e.g. protonated His, Gua, Trp)
	NB	sp ³ nitrogen in 5-membered ring with lone pair (e.g. N7 in purines)
	NC	sp ³ nitrogen in 6-membered ring with lone pair (e.g. N3 in purines)
	N*	sp ³ nitrogen in 5-membered ring with carbon substituent (in purine nucleotides)
	N2	sp ³ nitrogens of aromatic amines and guanidinium ions
	N3	sp ³ nitrogen
oxygen	OW	sp ³ oxygen in TIP3P water
	OH	sp ³ oxygen in alcohols, tyrosine, and protonated carboxylic acids
	OS	sp ³ oxygen in ethers
	O	sp ³ oxygen in amides
	O2	sp ³ oxygen in ascaric acids
	S	sulfur in methionine and cysteine
	SH	sulfur in cysteine
sulfur	P	phosphorus in phosphates
hydrogen	H	H attached to N
	HW	H in TIP3P water
	HO	H in alcohols and acids
	HS	H attached to sulfur
	HA	H attached to aromatic carbon
	HC	H attached to aliphatic carbon with no electron-withdrawing substituents
	H1	H attached to aliphatic carbon with one electron-withdrawing substituent
	H2	H attached to aliphatic carbon with two electron-withdrawing substituents
	H3	H attached to aliphatic carbon with three electron-withdrawing substituents
	HP	H attached to carbon directly bonded to formally positive atom (e.g. Cε next to NH ⁺ of lysine)
	H4	H attached to aromatic carbon with one electronegative neighbor (e.g. hydrogens on C5 of Trp, C6 of Thy)
	H5	H attached to aromatic carbon with two electronegative neighbors (e.g. H8 of Ade and H2 of Ade)

AMBER Atom Types

TYPES					
C	12.0100	1.9080	0.0860	HS	1.0080 0.6000 0.0157
C*	12.0100	1.9080	0.0860	HW	1.0080 0.0000 0.0000
CA	12.0100	1.9080	0.0860	HU	1.0080 0.0000 0.0000
CB	12.0100	1.9080	0.0860	HV	1.0080 0.0000 0.0000
CC	12.0100	1.9080	0.0860	N	14.0100 1.8240 0.1700
CM	12.0100	1.9080	0.0860	N*	14.0100 1.8240 0.1700
CN	12.0100	1.9080	0.0860	N2	14.0100 1.8240 0.1700
CV	12.0100	1.9080	0.0860	N3	14.0100 1.8240 0.1700
CW	12.0100	1.9080	0.0860	NA	14.0100 1.8240 0.1700
CR	12.0100	1.9080	0.0860	NB	14.0100 1.8240 0.1700
CT	12.0100	1.9080	0.1094	NC	14.0100 1.8240 0.1700
CV	12.0100	1.9080	0.0860	O	16.0000 1.6612 0.2100
CW	12.0100	1.9080	0.0860	O2	16.0000 1.6612 0.2100
H	1.0080	0.6000	0.0157	OH	16.0000 1.7210 0.2104
H1	1.0080	1.3870	0.0157	OS	16.0000 1.6837 0.1700
H2	1.0080	1.2870	0.0157	OW	16.0000 1.7683 0.1520
H3	1.0080	1.1870	0.0157	OT	16.0000 1.7683 0.1520
H4	1.0080	1.4090	0.0150	OJ	16.0000 1.7699 0.1550
H5	1.0080	1.3590	0.0150	MU	1.0000 0.0000 0.0000
HA	1.0080	1.4590	0.0150	OV	16.0000 1.7766 0.1554
HC	1.0080	1.4870	0.0157	P	30.9700 2.1000 0.2000
HO	1.0080	0.0000	0.0000	S	32.0600 2.0000 0.2500
HP	1.0080	1.1000	0.0157	SH	32.0600 2.0000 0.2500

AMBER Bond Parameters

$$\sum_{\text{bonds}} K_r (r - r_{eq})^2$$

BONDS	553.00	0.9572	CB-N*	436.00	1.3740	OB-P	230.00	1.6100
OW-HW	553.00	1.5139	CB-NB	414.00	1.3910	OB-P	230.00	1.6100
IW-HW	553.00	1.5139	CB-NC	461.00	1.3540	C*-HC	367.00	1.0800
OY-HW	553.00	0.9572	CB-HE	367.00	1.0800	C-N	490.00	1.2350
OU-HU	553.00	0.9572	CK-N*	440.00	1.3710	C*-CB	388.00	1.4590
HU-HU	553.00	1.5139	CK-NB	529.00	1.3040	C*-CT	317.00	1.4950
OU-HU	553.00	0.1500	CM-CN	549.00	1.2500	C*-CW	546.00	1.2520
OU-HV	553.00	0.8735	CM-CT	317.00	1.5100	CA-CN	469.00	1.4000
OU-HV	553.00	1.0000	CM-HA	367.00	1.0800	CA-CN	447.00	1.4190
HV-HV	553.00	1.6330	CM-H4	367.00	1.0800	CC-CT	317.00	1.5040
C-CA	469.00	1.4090	CM-HS	367.00	1.0800	CC-CT	317.00	1.5040
C-CB	447.00	1.4190	CM-N*	448.00	1.3650	CC-CW	518.00	1.3710
C-CH	410.00	1.4440	CO-HS	367.00	1.0800	CC-NA	422.00	1.3850
C-CT	317.00	1.5220	CO-NC	502.00	1.3240	CC-NB	410.00	1.3940
C-N*	424.00	1.3830	CT-CT	310.00	1.5260	CN-NA	428.00	1.3800
C-NA	418.00	1.3880	CT-HC	340.00	1.0900	CK-HS	367.00	1.0800
C-NC	457.00	1.2580	CT-H1	340.00	1.0900	CK-NA	477.00	1.2430
C-O	570.00	1.2290	CT-H2	340.00	1.0900	CK-NB	488.00	1.3350
C-O2	656.00	1.2500	CT-H3	340.00	1.0900	CT-N	337.00	1.4490
C-OH	450.00	1.2640	CT-H4	340.00	1.0900	CT-H3	367.00	1.4710
CA-CA	469.00	1.4000	CT-N*	337.00	1.4750	CT-S	227.00	1.8100
CA-CB	469.00	1.4040	CT-N2	337.00	1.4630	CT-SH	237.00	1.8100
CA-CM	427.00	1.4330	CT-OH	320.00	1.4100	CV-H4	367.00	1.0800
CA-CT	317.00	1.5100	CT-OS	320.00	1.4100	CV-NB	410.00	1.3940
CA-HA	367.00	1.0800	H-N2	434.00	1.0100	CW-H4	367.00	1.0800
CA-H4	367.00	1.0800	H-N3	434.00	1.0100	CW-NA	427.00	1.3810
CA-N2	481.00	1.3400	H-NA	434.00	1.0100	CV-NB	410.00	1.3940
CA-NA	427.00	1.3810	HO-OH	553.00	0.9600	H-N3	434.00	1.0100
CA-NC	483.00	1.3390	HO-OS	553.00	0.9600	HS-SH	274.00	1.2360
CB-CB	520.00	1.3700	O2-P	525.00	1.4800	S-P	166.00	2.0380

AMBER Bond Angle Parameters

$$\sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2$$

AMBER	100.00	104.52	C-C-NB	63.00	109.20	H1-CT-CT	90.00	109.20	C-NB-CT	30.00	119.50
OW-OW-HW	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
IW-OW-HW	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
OY-OW-HW	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
OU-HU	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
HU-HU	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
OU-HU	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
OU-HV	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
OU-HV	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
HV-HV	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-CA	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-CB	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-CH	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-CT	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-N*	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-NA	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-NC	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-O	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-O2	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
C-OH	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-CA	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-CB	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-CM	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-CT	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-HA	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-H4	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-N2	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-NA	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CA-NC	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50
CB-CB	100.00	104.52	C-C-NB	70.00	109.20	H1-CT-CT	90.00	114.70	C-NB-CT	30.00	119.50

AMBER Dihedral Angle Parameters

$$\sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$$

DIEHEDRALS									
X -X -C -O	4	14.500	180.00	2.00	X -CB-CH-X	4	12.000	180.00	2.00
X -O2-C -O2	4	12.000	180.00	2.00	X -CC-CT-X	4	0.000	0.00	2.00
X -X -N -H	4	8.700	180.00	2.00	X -CC-CV-X	4	20.600	180.00	2.00
X -O2-C -O2	4	5.800	180.00	2.00	X -CC-OM-X	4	21.500	180.00	2.00
X -O2-C -O2	4	5.400	180.00	2.00	X -CC-NA-X	4	5.600	180.00	2.00
X -O2-C -O2	4	8.000	180.00	2.00	X -CB-NB-X	4	4.800	180.00	2.00
X -O2-C -O2	2	1.800	180.00	2.00	X -CN-NA-X	4	6.100	180.00	2.00
X -O2-C -O2	4	14.500	180.00	2.00	X -CB-NB-X	2	10.000	180.00	2.00
X -O2-C -O2	4	14.000	180.00	2.00	X -CB-NB-X	2	4.800	180.00	2.00
X -O2-C -O2	4	10.200	180.00	2.00	X -CN-NA-X	4	6.000	180.00	2.00
X -O2-C -O2	6	0.000	0.00	3.00	CT-CT-OS-CT	1	0.983	180.00	-3.00
X -O2-C -O2	4	9.600	180.00	2.00	CT-CT-OS-CT	1	0.100	180.00	2.00
X -O2-C -O2	4	9.000	180.00	2.00	CT-N -CT-C	1	0.200	180.00	2.00
X -O2-C -O2	2	9.600	180.00	2.00	N -CT-C -N	1	0.400	180.00	-4.00
X -O2-C -O2	4	6.000	180.00	2.00	N -CT-N -H	1	0.750	180.00	1.00
X -O2-C -O2	4	21.600	180.00	2.00	CT-CT-N -C	1	0.500	180.00	-4.00
X -O2-C -O2	2	5.100	180.00	2.00	CT-CT-N -C	1	0.150	180.00	-3.00
X -O2-C -O2	2	8.300	180.00	2.00	CT-CT-N -C	1	0.530	0.00	1.00
X -O2-C -O2	4	6.800	180.00	2.00	CT-CT-C -N	1	0.100	0.00	-4.00
X -O2-C -O2	2	20.000	180.00	2.00	CT-CT-C -N	1	0.070	0.00	2.00
X -O2-C -O2	4	26.600	180.00	2.00	H -N -C -O	1	2.500	180.00	-2.00
X -O2-C -O2	6	0.000	0.00	3.00	H -N -C -O	1	2.000	0.00	1.00
X -O2-C -O2	4	7.400	180.00	2.00	CT-N -S -CT	1	0.600	0.00	3.00
X -O2-C -O2	2	13.600	180.00	2.00	OS-CT-CT-OS	1	3.500	0.00	-3.00
X -O2-C -O2	9	1.400	0.00	3.00	OS-CT-CT-OS	1	1.000	0.00	2.00
X -O2-C -O2	9	1.000	0.00	3.00	OS-CT-CT-OS	1	1.444	0.00	-3.00
X -O2-C -O2	6	0.000	0.00	2.00	OS-CT-CT-OH	1	1.000	0.00	2.00
X -O2-C -O2	3	1.150	0.00	3.00	OS-CT-CT-OH	1	0.444	0.00	-3.00
X -O2-C -O2	3	0.500	0.00	3.00	OS-CT-CT-OH	1	1.000	0.00	2.00
X -O2-C -O2	3	0.750	0.00	3.00	OS-CT-CT-OH	1	0.250	0.00	-3.00
X -O2-C -O2	3	0.750	0.00	3.00	OS-CT-CT-OH	1	1.000	0.00	2.00
X -O2-C -O2	3	0.750	0.00	3.00	OS-CT-CT-OH	1	0.250	0.00	-3.00
X -O2-C -O2	4	10.000	180.00	2.00	OS-P -OS-CT	1	1.200	0.00	2.00
X -O2-C -O2	9	1.400	0.00	3.00	OS-P -OS-CT	1	0.250	0.00	-3.00
X -O2-C -O2	9	1.000	0.00	3.00	OS-CT-N* -C	1	1.200	0.00	2.00
X -O2-C -O2	3	0.750	0.00	3.00	OS-CT-N* -C	1	0.500	180.00	-2.00
X -O2-C -O2	4	6.700	180.00	2.00	OS-CT-N* -C	1	2.500	0.00	1.00
X -O2-C -O2	6	0.000	0.00	2.00	OS-CT-N* -C	1	0.500	180.00	-2.00
X -O2-C -O2	4	26.100	180.00	2.00	OS-CT-N* -C	1	2.500	0.00	1.00
X -O2-C -O2	4	14.500	180.00	2.00	OS-CT-N* -C	1	2.500	0.00	1.00

AMBER Improper Dihedral Angle Parameters

$$\sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$$

IMPROPER DIHEDRALS									
X -X -C -O	10.50	180.00	2.00	X -X -CA-H4	1.10	180.00	2.00		
X -O2-C -O2	10.50	180.00	2.00	X -X -CA-H5	1.10	180.00	2.00		
X -X -N -H	1.00	180.00	2.00	CK-CB-N*-CT	1.00	180.00	2.00		
X -X -N2-H	1.00	180.00	2.00	CM-C -N*-CT	1.00	180.00	2.00		
X -X -NA-H	1.00	180.00	2.00	CM-C -CM-CT	1.10	180.00	2.00		
X -N2-CA-N2	10.50	180.00	2.00	NA-CV-CC-CT	1.10	180.00	2.00		
X -CT-N -CT	1.00	180.00	2.00	NB-CV-CC-CT	1.10	180.00	2.00		
X -X -CA-HA	1.10	180.00	2.00	NA-CW-CC-CT	1.10	180.00	2.00		
X -X -CW-H4	1.10	180.00	2.00	CW-CB-C*-CT	1.10	180.00	2.00		
X -X -CR-H5	1.10	180.00	2.00	CA-CA-CA-CT	1.10	180.00	2.00		
X -X -CV-H4	1.10	180.00	2.00	C -CM-CM-CT	1.10	180.00	2.00		
X -X -CQ-H5	1.10	180.00	2.00	NC-CM-CA-N2	1.10	180.00	2.00		
X -X -CK-H5	1.10	180.00	2.00	CB-NC-CA-N2	1.10	180.00	2.00		
X -X -CM-H4	1.10	180.00	2.00	NA-NC-CA-N2	1.10	180.00	2.00		
X -X -CM-HA	1.10	180.00	2.00	CA-CA-C -OH	1.10	180.00	2.00		

Nichtbindende Wechselwirkung

$$\sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

- Summe über alle Atompaaire (i,j), die nicht über 1 oder 2 kovalente Bindungen verbunden sind
- Lennard-Jones Potenzial + elektrostatische Wechselwirkung
- Funktion der interatomaren Abstände R_{ij}
- Parameter, von den Typen der Atome abhängig:
 - A_{ij}, B_{ij} Parameter des Lennard-Jones Potentials
 - q_i, q_j Partialladungen
 - ϵ Dielektrizitätskonstante
- Ohne besondere Vorkehrungen wächst der Rechenaufwand mit dem Quadrat der Molekülgröße
- Auch im Torsionswinkelraum vorhanden

Lennard-Jones Potenzial

$$\sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right]$$

- Schwacher anziehender Beitrag: Van der Waals Kraft
- Starke Abstoßung für kurze Abstände R_{ij} aufgrund des Pauli-Prinzips
- $A_{ij} = V_{j,\min} R_{j,\min}^{12}, B_{ij} = 2 V_{j,\min} R_{j,\min}^6$ mit
 - $V_{j,\min} = (V_{i,\min} V_{j,\min})^{1/2}$ Potenzialminimum
 - $R_{j,\min} = R_{i,\min} + R_{j,\min}$ energetisch günstigster Abstand
- Kurzreichweitig

AMBER Atom Types

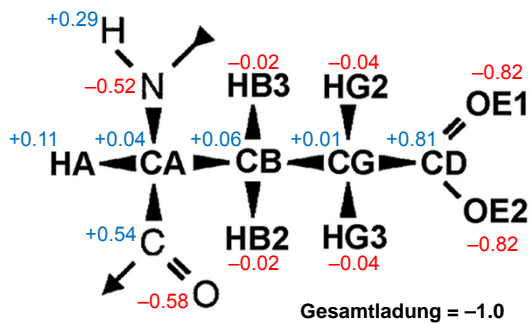
ATOM TYPE	MASS	$R_{i,\min}$	$V_{i,\min}$
C	12.0100	1.9080	0.0860
C*	12.0100	1.9080	0.0860
CA	12.0100	1.9080	0.0860
CB	12.0100	1.9080	0.0860
CC	12.0100	1.9080	0.0860
CK	12.0100	1.9080	0.0860
CM	12.0100	1.9080	0.0860
CN	12.0100	1.9080	0.0860
CQ	12.0100	1.9080	0.0860
CR	12.0100	1.9080	0.0860
CT	12.0100	1.9080	0.1094
CV	12.0100	1.9080	0.0860
CW	12.0100	1.9080	0.0860
H	1.0080	0.6000	0.0157
H1	1.0080	1.3870	0.0157
H2	1.0080	1.2870	0.0157
H3	1.0080	1.1870	0.0157
H4	1.0080	1.4090	0.0150
H5	1.0080	1.3590	0.0150
HA	1.0080	1.4590	0.0150
HA	1.0080	1.4870	0.0157
HO	1.0080	0.0000	0.0000
HP	1.0080	1.1000	0.0157
HS	1.0080	1.0080	0.0157
HW	1.0080	0.0000	0.0000
HU	1.0080	0.0000	0.0000
HV	1.0080	0.0000	0.0000
N	14.0100	1.8240	0.1700
N*	14.0100	1.8240	0.1700
N2	14.0100	1.8240	0.1700
N3	14.0100	1.8240	0.1700
NA	14.0100	1.8240	0.1700
NB	14.0100	1.8240	0.1700
NC	14.0100	1.8240	0.1700
O	16.0000	1.6612	0.2100
O2	16.0000	1.6612	0.2100
OH	16.0000	1.7210	0.2104
OS	16.0000	1.6837	0.1700
OW	16.0000	1.7683	0.1520
OT	16.0000	1.7683	0.1520
OU	16.0000	1.7699	0.1550
MU	1.0000	0.0000	0.0000
OV	16.0000	1.7766	0.1554
P	30.9700	2.1000	0.2000
S	32.0600	2.0000	0.2500
SH	32.0600	2.0000	0.2500

Elektrostatisches (Coulomb) Potenzial

$$\sum_{i < j} \frac{q_i q_j}{\epsilon R_{ij}}$$

- Partialladungen q_i, q_j
- Dielektrizitätskonstante ϵ beschreibt (gemittelte) Polarisierbarkeit der Umgebung
 - $\epsilon_{\text{Wasser}} \approx 80; \epsilon_{\text{apolar}} \approx 2.$
 - Werden alle Ladungen explizit behandelt: $\epsilon = 1$
- Langreichweitig!
- Stark!
- Wichtigster Energiebeitrag auch für Wasserstoffbrücken

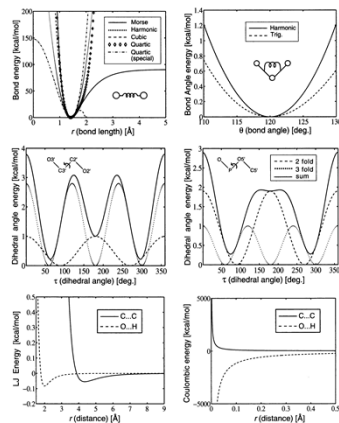
Partiellladungen (Glu, AMBER)



Bestimmung von Kraftfeldparametern

- **Fit der Energie zu quantenmechanischen *ab initio* Rechnungen**
 - z.B. Partiellladungen
- **Fit zu experimentellen Daten**
 - Infrarot-Spektroskopie (Vibrationen)
 - Röntgenkristallographie (Geometrie)
 - NMR Spektroskopie
 - Thermodynamische Eigenschaften
- **Schwierigkeiten**
 - Exakte Beziehung molekulare Eigenschaft \leftrightarrow Kraftfeldparameter
 - Konzeptionelle Einfachheit \leftrightarrow Anzahl Parameter
 - Gegenseitige Abhängigkeit von Kraftfeldparametern
 - Widersprüchliche Anforderungen für Verbesserungen der Parameter
 - Einfluss von Näherungen
- **Genauigkeit wird durch den ungenaueren Term der Energiefunktion bestimmt.**

Energierterme



Kraftberechnung

- Kraft = -Gradient der potentiellen Energie
- Analytisch berechnete partielle Ableitungen der potentiellen Energie nach den Koordinaten aller Atome sind notwendig für klassische MD Simulation.
- MD im Torsionswinkelraum benötigt partielle Ableitungen der potentiellen Energie nach den Torsionswinkeln.

Paarlisten

- N Atome \rightarrow ca. $N(N-1)/2$ Paarwechselwirkungen
- Berechnung der nichtbindenden Wechselwirkung ist aufwendigster Teil der Kraftberechnung
- Potenzial nimmt mit der Entfernung ab \rightarrow Vernachlässigung von kleinen Termen = langen Distanzen
- Kurze Distanzen $R < R_{\max}$ werden periodisch in Paarliste gespeichert
- Wechselwirkungen werden nur für Atumpaare mit Abstand $R < R_{\text{cutoff}} < R_{\max}$ berechnet
- Langreichweitige elektrostatische WW $\rightarrow R_{\text{cutoff}} \approx 12 \text{ \AA}$
- Aktualisierung der Paarliste nach einer festen Anzahl Zeitschritten oder wenn sich ein Teilchen $R_{\max} - R_{\text{cutoff}}$ weit bewegt hat.

Paarliste und Cutoff

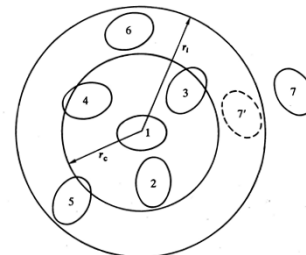
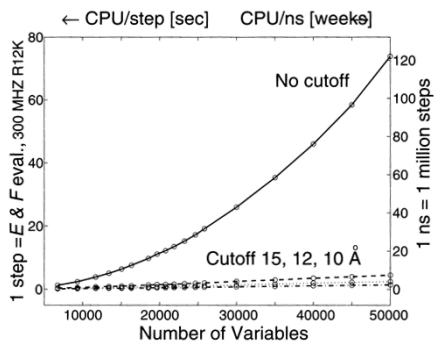


Fig. 5.4 The cutoff sphere, and its skin, around a molecule 1. Molecules 2, 3, 4, 5, and 6 are on the list of molecule 1; molecule 7 is not. Only molecules 2, 3, and 4 are within the range of the potential at the time the list is constructed.

Rechenzeit mit und ohne Cutoff



MD Algorithmen

Strukturelle Modellierung
(Masterstudiengang Bioinformatik)

MD Algorithmen

Sommersemester 2012

Peter Güntert

MD Algorithmen

- Energieminimierung
- Integration der Bewegungsgleichungen
- Temperaturkontrolle
- Druckkontrolle
- Periodische Randbedingungen

Energieflächen

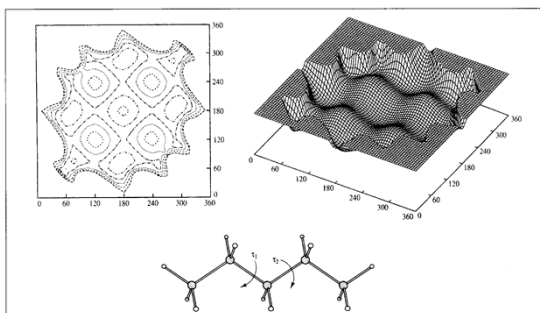


Fig. 5.1: Variation in the energy of pentane with the two torsion angles indicated and represented as a contour diagram and isometric plot. Only the lowest-energy regions are shown.

Lokale Minima, globales Minimum

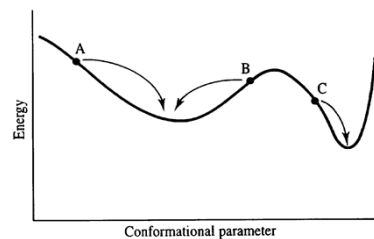


Fig. 5.3: A schematic one-dimensional energy surface. Minimisation methods move downhill to the nearest minimum. The statistical weight of the narrow, deep minimum may be less than a broad minimum which is higher in energy.

Energieminimierungsalgorithmen

- Ohne Ableitungen
- Mit Gradienten:
 - Steilster Abstieg (steepest descent)
 - konjugierte Gradienten (conjugate gradients)
- Mit zweiter Ableitung: Newton-Raphson Methode

Simplex Algorithmus

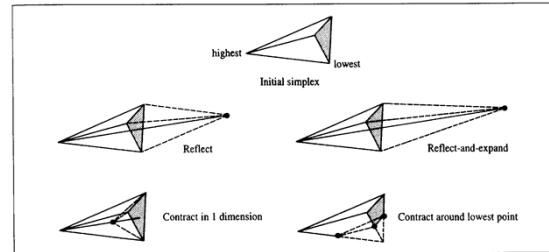


Fig. 5.4: The three basic moves permitted to the simplex algorithm (reflection, and its close relation reflect-and-expand; contract in one dimension and contract around the lowest point). (Figure adapted from Press W H, B P Flannery, S A Teukolsky and W T Vetterling 1992. Numerical Recipes in Fortran. Cambridge, Cambridge University Press.)

Simplex Algorithmus

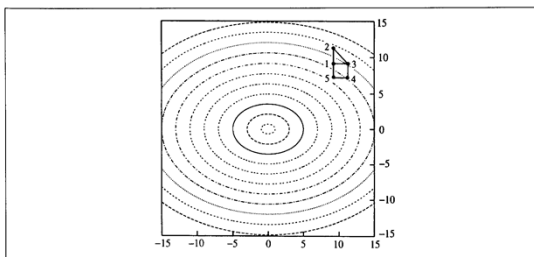


Fig. 5.5: The first few steps of the simplex algorithm with the function $x^2 + 2y^2$. The initial simplex corresponds to the triangle 123. Point 2 has the largest value of the function and the next simplex is the triangle 134. The simplex for the third step is 145.

Eindimensionale Minimierung

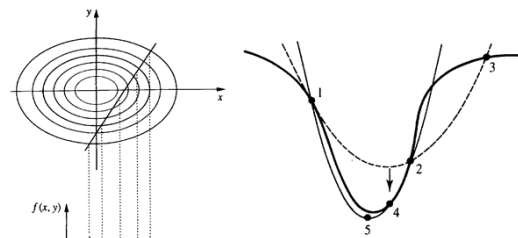


Fig. 5.7: A line search is used to locate the minimum in the function in the direction of the gradient.

Fig. 5.8: The minimum in a line search may be found more effectively by fitting an analytical function such as a quadratic to the initial set of three points (1,2 and 3). A better estimate of the minimum can then be found by fitting a new function to the points 1, 2 and 4 and finding its minimum.

Energieminimierung: Steepest descent

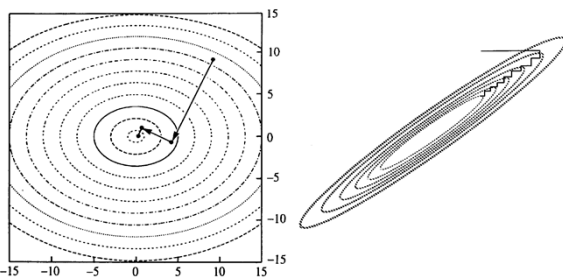


Fig. 5.9: Application of steepest descents to the function $x^2 + 2y^2$.

Fig. 5.10: The steepest descents method can give undesirable behaviour in a long narrow valley.

Konjugierte Gradientenmethode

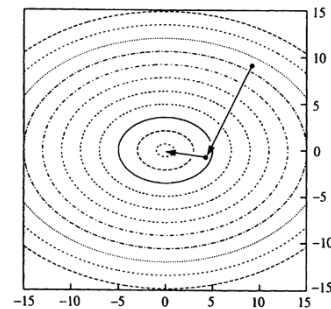


Fig. 5.11: Application of conjugate gradients method to the function $x^2 + 2y^2$.

Taylor Reihenentwicklung

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f'''(x_0)}{3!}(x - x_0)^3 + \dots$$

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \frac{1}{6} \delta t^3 \mathbf{b}(t) + \frac{1}{24} \delta t^4 \mathbf{c}(t) + \dots$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \delta t \mathbf{a}(t) + \frac{1}{2} \delta t^2 \mathbf{b}(t) + \frac{1}{6} \delta t^3 \mathbf{c}(t) + \dots$$

$$\mathbf{a}(t + \delta t) = \mathbf{a}(t) + \delta t \mathbf{b}(t) + \frac{1}{2} \delta t^2 \mathbf{c}(t) + \dots$$

$$\mathbf{b}(t + \delta t) = \mathbf{b}(t) + \delta t \mathbf{c}(t) + \dots$$

Verlet Algorithmus

The Verlet algorithm uses the positions and accelerations at time t , and the positions from the previous step, $\mathbf{r}(t - \delta t)$, to calculate the new positions at $t + \delta t$, $\mathbf{r}(t + \delta t)$. We can write down the following relationships between these quantities and the velocities at time t :

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \dots \tag{7.6}$$

$$\mathbf{r}(t - \delta t) = \mathbf{r}(t) - \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) - \dots \tag{7.7}$$

Adding these two equations gives

$$\mathbf{r}(t + \delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \delta t) + \delta t^2 \mathbf{a}(t) \tag{7.8}$$

The velocities do not explicitly appear in the Verlet integration algorithm. The velocities can be calculated in a variety of ways; a simple approach is to divide the difference in positions at times $t + \delta t$ and $t - \delta t$ by $2\delta t$:

$$\mathbf{v}(t) = [\mathbf{r}(t + \delta t) - \mathbf{r}(t - \delta t)] / 2\delta t \tag{7.9}$$

Alternatively, the velocities can be estimated at the half-step, $t + \frac{1}{2} \delta t$:

$$\mathbf{v}(t + \frac{1}{2} \delta t) = [\mathbf{r}(t + \delta t) - \mathbf{r}(t)] / \delta t \tag{7.10}$$

Leap-frog Algorithmus

The *leap-frog* algorithm uses the following relationships:

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t + \frac{1}{2} \delta t) \tag{7.11}$$

$$\mathbf{v}(t + \frac{1}{2} \delta t) = \mathbf{v}(t - \frac{1}{2} \delta t) + \delta t \mathbf{a}(t) \tag{7.12}$$

To implement the leap-frog algorithm, the velocities $\mathbf{v}(t + \frac{1}{2} \delta t)$ are first calculated from the velocities at time $t - \frac{1}{2} \delta t$ and the accelerations at time t . The positions $\mathbf{r}(t + \delta t)$ are then deduced from the velocities just calculated together with the positions at time $\mathbf{r}(t)$ using Equation (7.11). The velocities at time t can be calculated from

$$\mathbf{v}(t) = \frac{1}{2} [\mathbf{v}(t + \frac{1}{2} \delta t) + \mathbf{v}(t - \frac{1}{2} \delta t)] \tag{7.13}$$

The velocities thus 'leap-frog' over the positions to give their values at $t + \frac{1}{2} \delta t$ (hence the name). The positions then leap over the velocities to give their new values at $t + \delta t$, ready for the velocities at $t + \frac{3}{2} \delta t$, and so on. The leap-frog method has two advantages over the

Velocity Verlet Algorithmus

The *velocity Verlet* method [Swope *et al.* 1982] gives positions, velocities and accelerations at the same time and does not compromise precision:

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) \tag{7.14}$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{1}{2} \delta t [\mathbf{a}(t) + \mathbf{a}(t + \delta t)] \tag{7.15}$$

The velocity Verlet method is actually implemented as a three-stage procedure because, as can be seen from Equation (7.15), to calculate the new velocities requires the accelerations at both t and $t + \delta t$. Thus in the first step the positions at $t + \delta t$ are calculated according to Equation (7.14) using the velocities and the accelerations at time t . The velocities at time $t + \frac{1}{2} \delta t$ are then determined using:

$$\mathbf{v}(t + \frac{1}{2} \delta t) = \mathbf{v}(t) + \frac{1}{2} \delta t \mathbf{a}(t) \tag{7.16}$$

New forces are next computed from the current positions, thus giving $\mathbf{a}(t + \delta t)$. In the final step, the velocities at time $t + \delta t$ are determined using:

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t + \frac{1}{2} \delta t) + \frac{1}{2} \delta t \mathbf{a}(t + \delta t) \tag{7.17}$$

Beemans Algorithmus

Beeman's algorithm [Beeman 1976] is also related to the Verlet method:

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{3}{8} \delta t^2 \mathbf{a}(t) - \frac{1}{8} \delta t^2 \mathbf{a}(t - \delta t) \tag{7.18}$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{3}{4} \delta t \mathbf{a}(t) + \frac{5}{8} \delta t \mathbf{a}(t) - \frac{1}{8} \delta t \mathbf{a}(t - \delta t) \tag{7.19}$$

The Beeman integration scheme uses a more accurate expression for the velocity. As a consequence it often gives better energy conservation, because the kinetic energy is calculated directly from the velocities. However, the expressions used are more complex than those of the Verlet algorithm and so it is computationally more expensive.

Gear Predictor-Corrector Algorithmen

The predictor-corrector methods [Gear 1971] form a general family of integration algorithms from which one can select a scheme that is correct to a given order. These methods have three basic steps. First, new positions, velocities, accelerations and higher-order terms are predicted according to the Taylor expansion, Equations (7.2)-(7.4). In the second stage, the forces are evaluated at the new positions to give accelerations $\mathbf{a}(t + \delta t)$. These accelerations are then compared with the accelerations that are predicted from the Taylor series expansion, $\mathbf{a}^p(t + \delta t)$. The difference between the predicted and calculated accelerations is then used to 'correct' the positions, velocities, etc., in the correction step:

$$\Delta \mathbf{a}(t + \delta t) = \mathbf{a}^c(t + \delta t) - \mathbf{a}(t + \delta t) \tag{7.22}$$

Then

$$\mathbf{r}^c(t + \delta t) = \mathbf{r}(t + \delta t) + c_0 \Delta \mathbf{a}(t + \delta t) \tag{7.23}$$

$$\mathbf{v}^c(t + \delta t) = \mathbf{v}(t + \delta t) + c_1 \Delta \mathbf{a}(t + \delta t) \tag{7.24}$$

$$\mathbf{a}^c(t + \delta t) / 2 = \mathbf{a}(t + \delta t) / 2 + c_2 \Delta \mathbf{a}(t + \delta t) \tag{7.25}$$

$$\mathbf{b}^c(t + \delta t) / 6 = \mathbf{b}(t + \delta t) / 6 + c_3 \Delta \mathbf{a}(t + \delta t) \tag{7.26}$$

Gear has suggested 'best' values of the coefficients c_0, c_1, \dots . The set of coefficients to use depends upon the order of the Taylor series expansion. In Equations (7.23)-(7.26) the expansion has been truncated after the third derivative of the positions (i.e. $\mathbf{b}(t)$). The appropriate set of coefficients to use in this case is $c_0 = \frac{1}{6}, c_1 = \frac{5}{6}, c_2 = 1$ and $c_3 = \frac{1}{3}$.

Energie-fluktuation

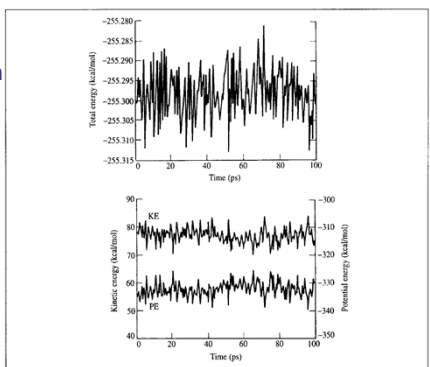


Fig. 7.3: Variation in total energy versus time for the production phase of a molecular dynamics simulation of 256 argon atoms at a temperature of 100 K and a density of 1.306 g cm⁻³ (top). The time step was 10 fs and the equations of motion were integrated using the velocity Verlet algorithm. The variations in the kinetic and potential energies are also shown (bottom). The graphs have different scales.

Zeitschrittlänge

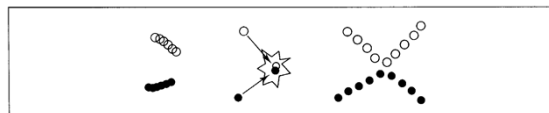


Fig. 7.4: With a very small time step (left) phase space is covered very slowly; a large time step (middle) gives instabilities. With an appropriate time step (right) phase space is covered efficiently and collisions occur smoothly.

System	Types of motion present	Suggested time step (s)
Atoms	Translation	10 ⁻¹⁴
Rigid molecules	Translation and rotation	5 × 10 ⁻¹⁵
Flexible molecules, rigid bonds	Translation, rotation, torsion	2 × 10 ⁻¹⁵
Flexible molecules, flexible bonds	Translation, rotation, torsion, vibration	10 ⁻¹⁵ or 5 × 10 ⁻¹⁶

Table 7.1 The different types of motion present in various systems together with suggested time steps.

Energieerhaltung

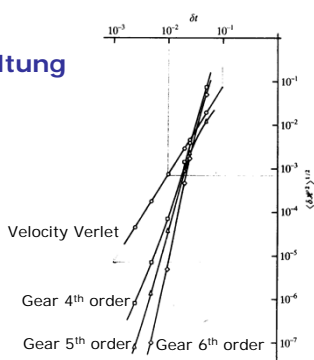


Fig. 3.3 Energy conservation of various algorithms. The system studied is as for Fig. 3.1. We calculate RMS energy fluctuations $(\Delta E^2)^{1/2}$ for various runs starting from the same initial condition, and proceeding for the same total simulation time t_{sim} , but using different time steps Δt and corresponding numbers of steps $t_{sim} = t_{step} \Delta t$. The plot uses log-log scales. The curves correspond to velocity Verlet (circles), Gear fourth-order (squares), Gear fifth-order (triangles), and Gear sixth-order (diamonds) algorithms.

Temperatur

Momentane Temperatur $T(t)$:

$$\frac{1}{2} N k_B T(t) = E_{kin}(t) = \sum_{i=1}^n \frac{1}{2} m_i v_i^2$$

N = Anzahl Freiheitsgrade ($N = 3n$), n = Anzahl Atome

Methoden für MD Simulation bei konstanter Temperatur:

- (strikt) konstante kinetische Energie und Temperatur
- erweitertes System mit zusätzlichem Freiheitsgrad
- schwache Kopplung an ein Wärmebad

Druck

Druck = Kraft pro Flächeneinheit auf die Wände des Systems
Aber: bei periodischen Randbedingungen keine Wand vorhanden

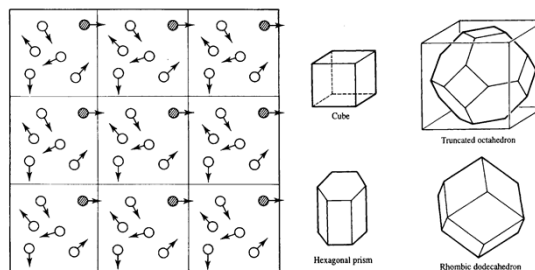
Virialsatz:
$$P = \frac{2}{3V} \left[E_{kin} + \frac{1}{2} \sum_{i=1}^n \vec{r}_{ij} \cdot \vec{F}_{ij} \right]$$

P = Druck, V = Volumen

Methoden für MD Simulation bei konstantem Druck:

- (strikt) konstanter Druck
- erweitertes System mit zusätzlichem Freiheitsgrad
- schwache Kopplung

Periodische Randbedingungen

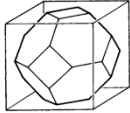


Periodische Randbedingungen

Rectangular box, side 2a (x) by 2b (y) by 2c (z)



Truncated octahedron derived from cube of side 2a



Hexagonal prism of length 2a (in z direction) and distance between opposite faces of the hexagon 2b



```

x = x - 2 * a * AINT(x/a)
y = y - 2 * b * AINT(y/b)
z = z - 2 * c * AINT(z/c)
A common alternative is:
x = x - a * ANINT(x/a)
y = y - b * ANINT(y/b)
z = z - c * ANINT(z/c)

x = x - 2 * a * AINT(x/a)
y = y - 2 * b * AINT(y/a)
z = z - 2 * c * AINT(z/a)
if (ABS(x) + ABS(y) + ABS(z)) >= 1.5 * A
then
  x = x - SIGN(a, x)
  y = y - SIGN(a, y)
  z = z - SIGN(a, z)
endif

z = z - 2 * a * AINT(z/a)
x = x - 2 * b * AINT(x/b)
if (ABS(x) + sqrt(3) * ABS(y)) >= 2 * b then
  x = x - SIGN(b, x)
  y = y - SIGN(sqrt(3) * b, y)
endif
    
```

Zeit- und Ensemblemittel

- Experimentelle Messgrößen sind Mittel über Raum und Zeit.
- Zeitmittel einer Messgröße A, die von den Positionen r und Impulsen p der N Teilchen im System abhängt:

$$A_{ave} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(r(t), p(t)) dt$$

- Ensemblemittel:

$$\langle A \rangle = \frac{1}{Z} \iint A(r, p) e^{-\frac{E(r, p)}{k_B T}} d^N r d^N p$$

$Z = \iint e^{-\frac{E(r, p)}{k_B T}} d^N r d^N p$ Zustandssumme
 $E(r, p)$ Gesamtenergie des Systems
 $k_B = 1.38 \cdot 10^{-23} J/K$ Boltzmannkonstante
 T absolute Temperatur

- Berechnung durch Mittelung über die Zeitschritte t_1, \dots, t_M einer MD Trajektorie:

$$\langle A \rangle \approx \frac{1}{M} \sum_{i=1}^M A(r(t_i), p(t_i))$$

Zeit- und Ensemblemittel

Experimentelle Messgrößen sind Mittel über Raum und Zeit

ü ä ö ß

MD Simulation im Torsionswinkelraum

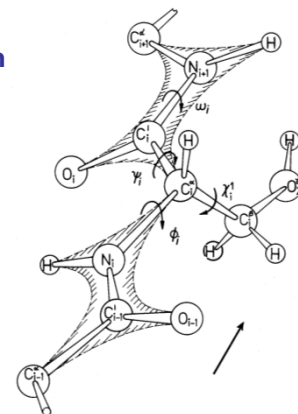
Strukturelle Modellierung
(Masterstudiengang Bioinformatik)

MD Simulation im Torsionswinkelraum

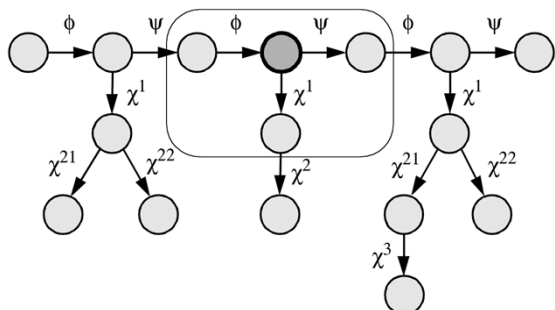
Sommersemester 2012

Peter Güntert

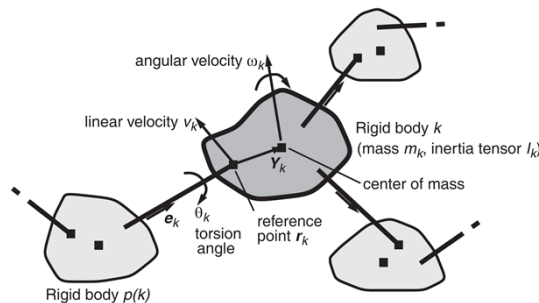
Diederwinkel in Polypeptiden



Tree structure of torsion angles

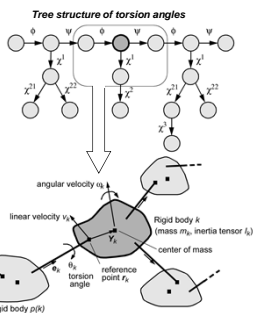


Tree structure with rigid bodies



Torsion angle dynamics algorithm

- Classical mechanical equations of motion are solved for a system with torsion angles as the only degrees of freedom.
- About 10 times less degrees of freedom than in conventional Cartesian space MD.
- Fixed bond lengths and bond angles:
 - no high frequency motions
 - longer integration time-steps
- Efficient algorithm required!
- Amount of computation proportional to system size



Programs DYANA, CYANA

Jain, Vaidehi, Rodriguez, *J. Comp. Phys.* 106, 258–268 (1993)
 Günter, Mumenthaler, Wüthrich, *J. Mol. Biol.* 273, 283–298 (1997)

Molecular Dynamics

Cartesian space		Torsion angle space
$E_{kin} = \frac{1}{2} \sum_{i=1}^N m_i \dot{x}_i^2$	Kinetic energy	$E_{kin} = \frac{1}{2} \sum_{k,l=1}^n M(\theta)_{kl} \dot{\theta}_k \dot{\theta}_l$
diagonal, constant (elements m_i)	Mass matrix M	non-diagonal, non-constant, $n \times n$
$\ddot{x}_i = \frac{1}{m_i} \frac{\partial E_{pot}}{\partial x_i}$	Accelerations	$M(\theta)\ddot{\theta} = C(\theta, \dot{\theta})$ (n linear equations)
proportional to N	Computational complexity	solving linear system of equations: $\sim n^3$ exploiting tree structure of the molecule: $\sim n$

Computation of the kinetic energy

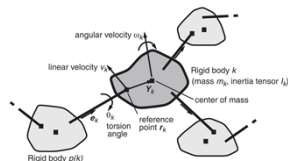
The angular velocity vector ω_k and the linear velocity v_k of the reference point of the rigid body k are calculated recursively from the corresponding quantities of the preceding rigid body $p(k)$:

$$\omega_k = \omega_{p(k)} + e_k \dot{\theta}_k$$

$$v_k = v_{p(k)} - (r_k - r_{p(k)}) \wedge \omega_{p(k)}$$

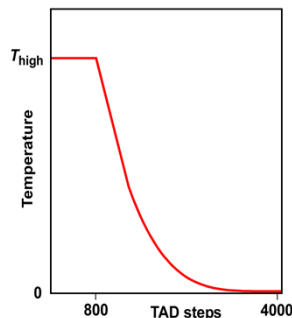
The kinetic energy can then be computed in a linear loop over all rigid bodies:

$$E_{kin} = \frac{1}{2} \sum_{k=0}^n [m_k v_k^2 + \omega_k \cdot I_k \omega_k + 2v_k \cdot (\omega_k \wedge m_k Y_k)] \quad [1]$$



Simulated annealing protocol

- Start from random structure
- Use all restraints simultaneously
- Adjustable parameters:
 - start temperature, T_{high}
 - number of TAD steps



Temperature control

Weak coupling to a heat bath is used to control the temperature:

$$\dot{\theta} \leftarrow \dot{\theta} \sqrt{1 + \frac{T^{\text{ref}} - T}{\tau T}}$$

$\dot{\theta}$ torsional velocities

T instantaneous temperature, $T = \frac{2E_{\text{kin}}}{nk_B}$
coupling constant

(Berendsen et al., J. Chem. Phys. 81, 3684–3690, 1984)

Time-step adaption

The same idea is used to adapt the integration time-step, Δt , such as to maintain a user-defined accuracy of energy conservation:

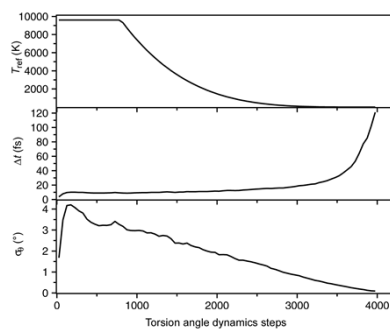
$$\Delta t \leftarrow \Delta t \sqrt{1 + \frac{\varepsilon^{\text{ref}} - \varepsilon}{\tau \varepsilon}}$$

ε is the relative change of the total energy, $E = E_{\text{kin}} + E_{\text{pot}}$:

$$\varepsilon = \left| \frac{E(t - \Delta t) - E(t)}{E(t)} \right|$$

Standard protocol: $\varepsilon^{\text{ref}} = 0.005 \dots 0.0001$, $\tau = 20$

Simulated annealing with torsion angle dynamics



CYANA Computation Time

