

Rechnen auf Grafikkarten

Christian Wiebeler

Inhaltsverzeichnis

- ▶ Einleitung
 - Geschichte des Spiels
 - Entwicklung Videospiele
 - Bedeutung der Grafikkarte
- ▶ Hauptteil
 - Anwendungen in der Teilchenphysik
 - QC auf GPU
- ▶ Schlussteil
 - Kommerzielle Verwendung in Forschung (QC)

Inhaltsverzeichnis

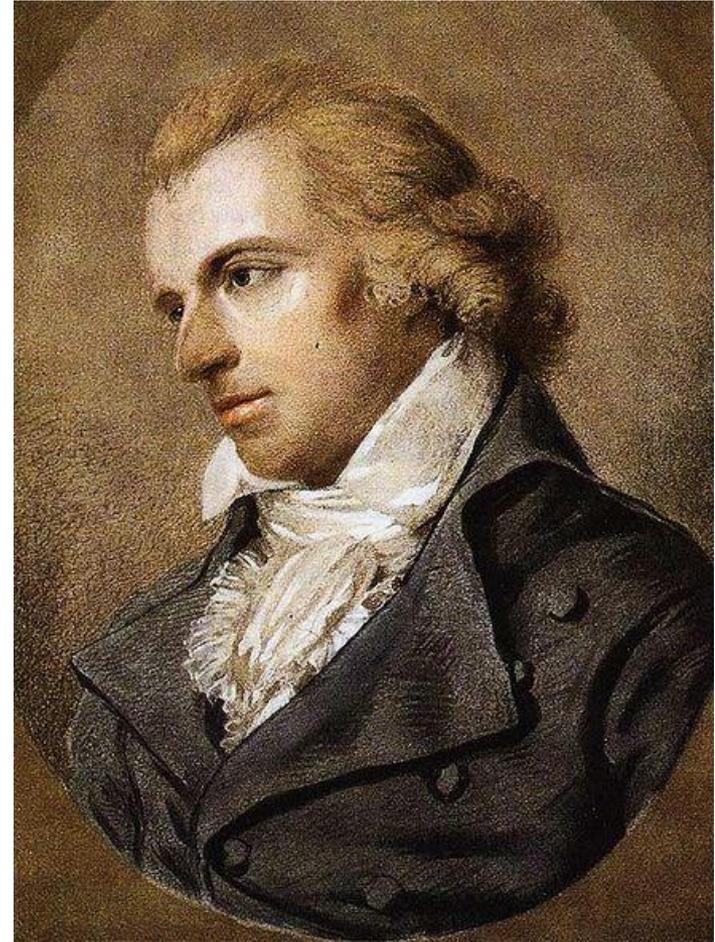
- ▶ **Einleitung**
 - Geschichte des Spiels
 - Entwicklung Videospiele
 - Bedeutung der Grafikkarte
- ▶ **Hauptteil**
 - Anwendungen in der Teilchenphysik
 - QC auf GPU
- ▶ **Schluss**
 - Kommerzielle Verwendung in Forschung (QC)

Panem et Circenses



Über die ästhetische Erziehung des Menschen

- ▶ „Der Mensch spielt nur, wo er in voller Bedeutung des Wortes Mensch ist, und er ist nur da ganz Mensch, wo er spielt.“



Homo ludens

- ▶ Wichtigkeit des Spielbegriffes für die Kultur bzw. allgemein für das Leben des Menschen in einer Gemeinschaft

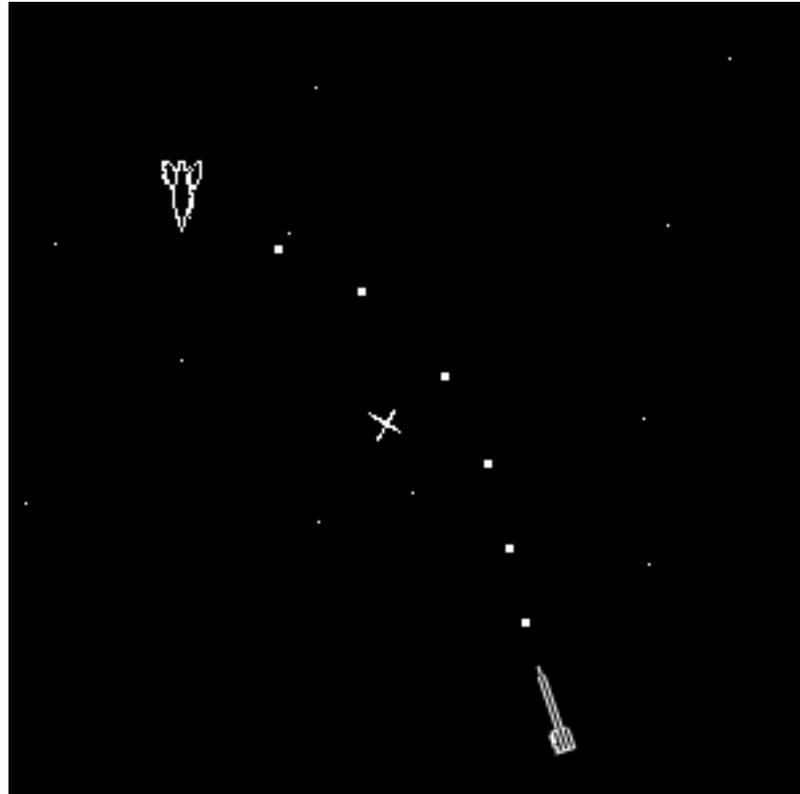


Spieltheorie

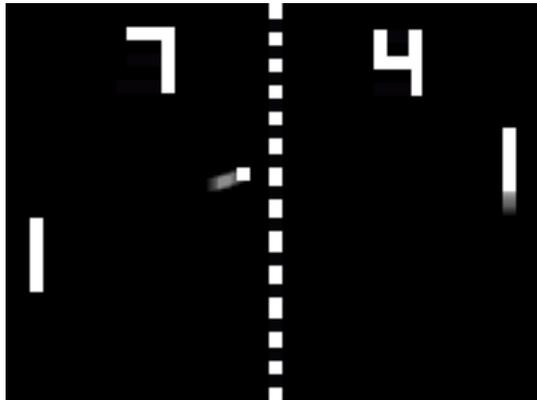
- ▶ Schönheitswettbewerb (Keynes)
- ▶ Spekulationsblasen
- ▶ Elfmeterschießen
- ▶ Lotto



Anfänge Videospiele



Die 1970er



PC und Konsolen



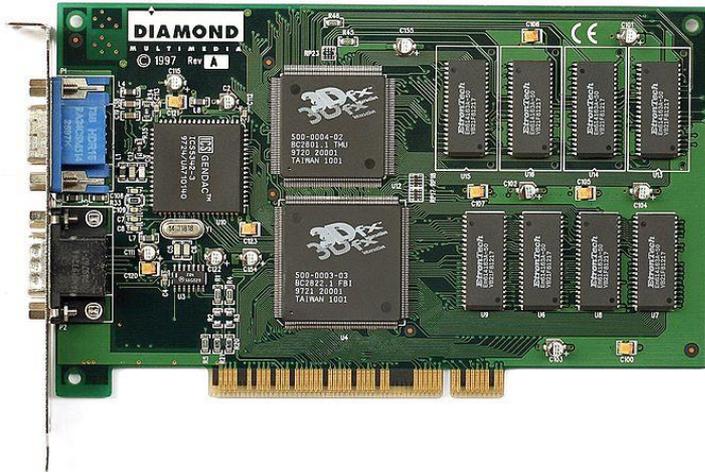
Neue Impulse aus Japan



Nintendo vs. Sega



3D Grafik



Entwicklung Spielegrafik



Aktueller Stand



Technische Daten

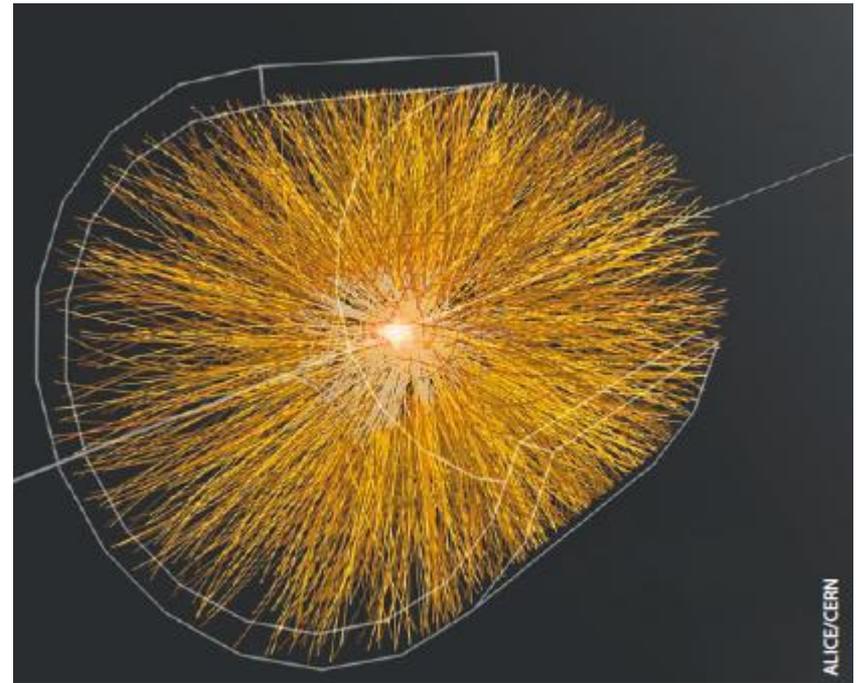
	Rechenleistung	Speicherbus-Datenrate
ATI Radeon HD 5870	2720,0 GFlops	153,6 GByte/s
NVIDIA GeForce GTX 295	1788,5 GFlops	223,8 GByte/s
Intel Core i7-970	94,0 GFlops	4,8 GByte/s
Intel Pentium 4 mit SSE3, 3,6 GHz	14,4 GFlops	5,0 GByte/s

Inhaltsverzeichnis

- ▶ **Einleitung**
 - Geschichte des Spiels
 - Entwicklung Videospiele
 - Bedeutung der Grafikkarte
- ▶ **Hauptteil**
 - Anwendungen in der Teilchenphysik
 - QC auf GPU
- ▶ **Schluss**
 - Kommerzielle Verwendung in Forschung (QC)

Teilchenphysik – Grafikkarten für die Datenflut

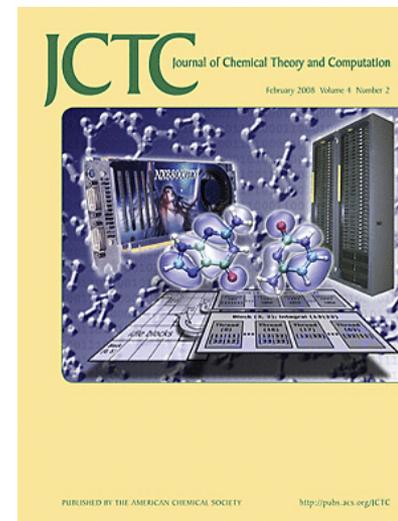
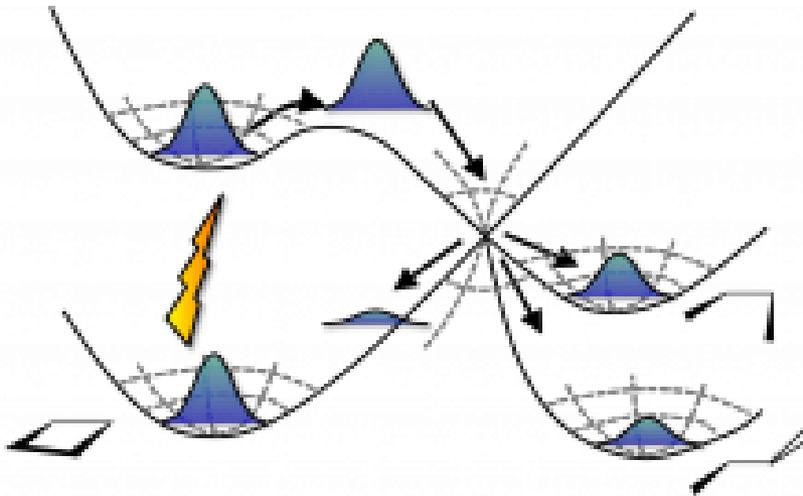
- ▶ Komplexere Detektorsysteme mit mehr Sensoren sind immer schneller auszulesen
- ▶ ALICE am LHC: Bis zu 20 Terabyte pro Sekunde
- ▶ Auswertung der Spur → Teilchenart



Dynamics of Ground and Excited Electronic States from First Principles

Plenar Vortrag – Todd J. Martinez

Frühjahrstagung der DPG in Dresden 2011



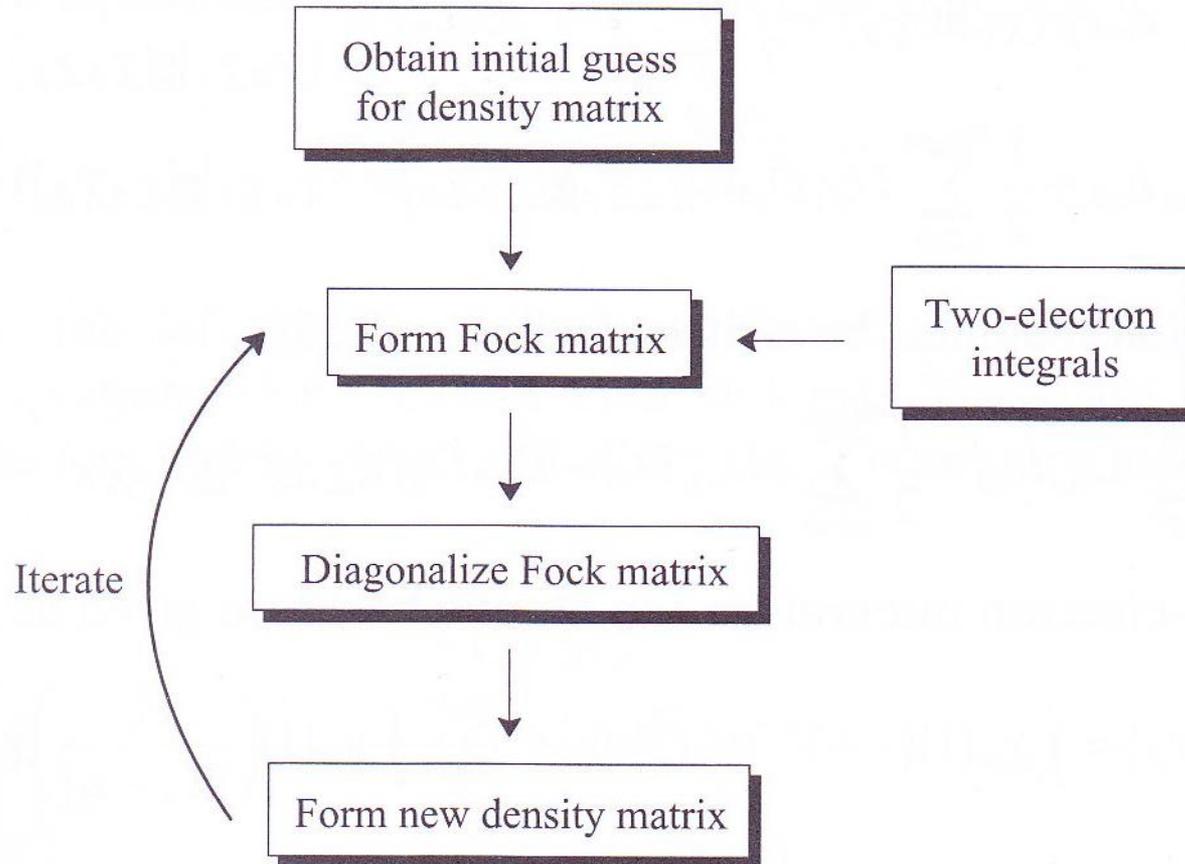
Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry. If mathematical analysis should ever hold a prominent place in chemistry – an aberration which is happily almost impossible – it would occasion a rapid and widespread degeneration of that science.

➤➤ Auguste Comte 1830

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

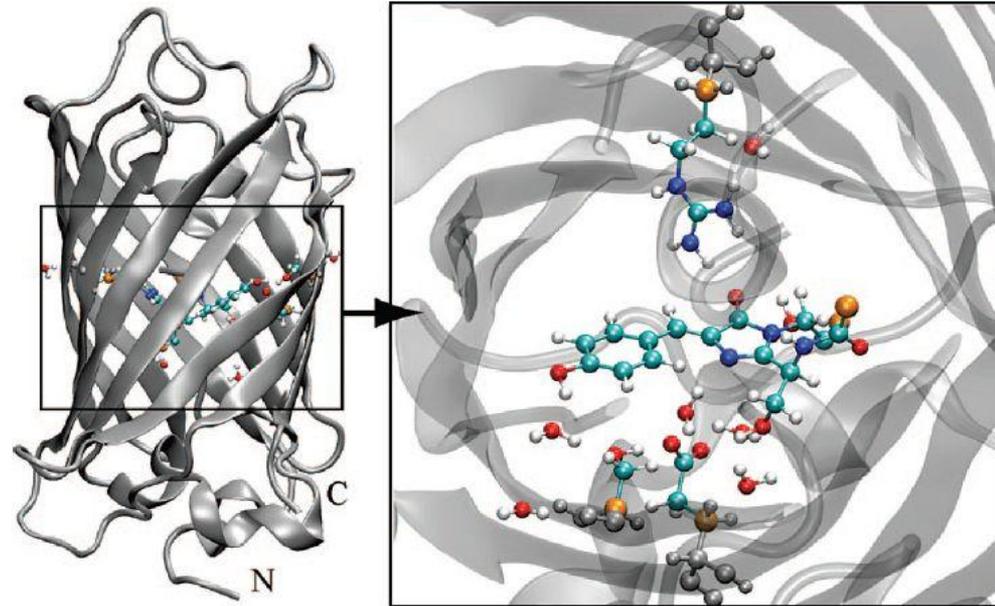
»» P. A. M. Dirac 1929

HF Algorithmus

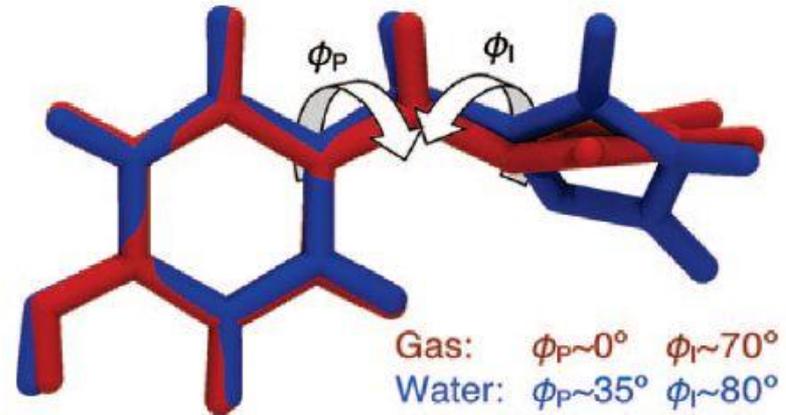
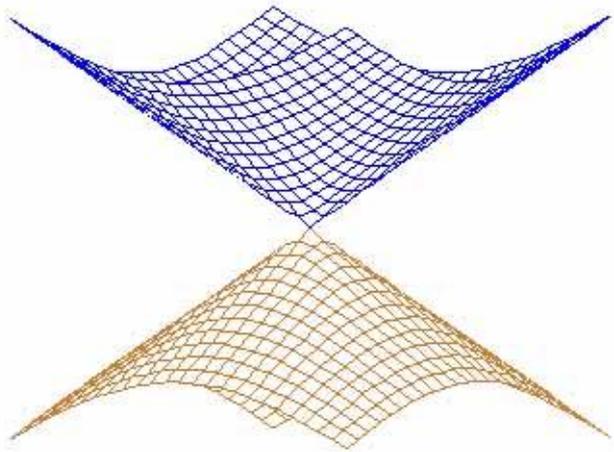


Modell & Methoden

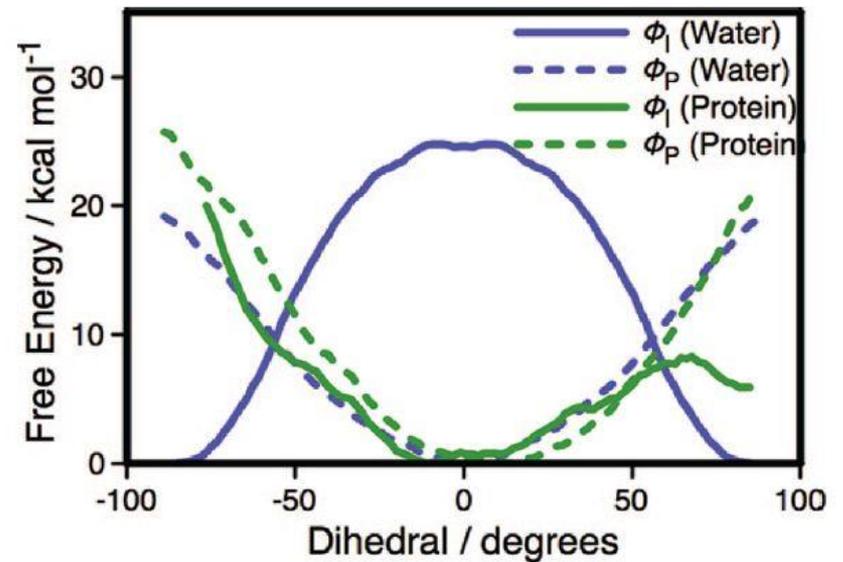
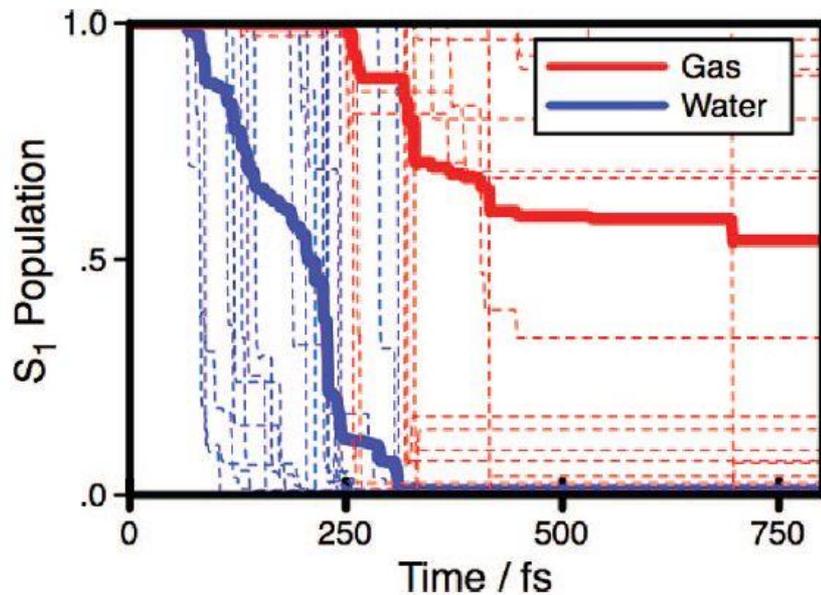
- Basissatz: FMS
- MM: Amber94
- QM:
 - Ab initio: SA-CASSCF (MRCI)
- TDDFT:
 - Nur angeregte Zustände mit Charakter einer einzigen Anregung zum Grundzustand
 - Scheitert bei der Beschreibung von Regionen nahe Conical Intersections
 - Für Photochemische Anwendungen nur bedingt geeignet.



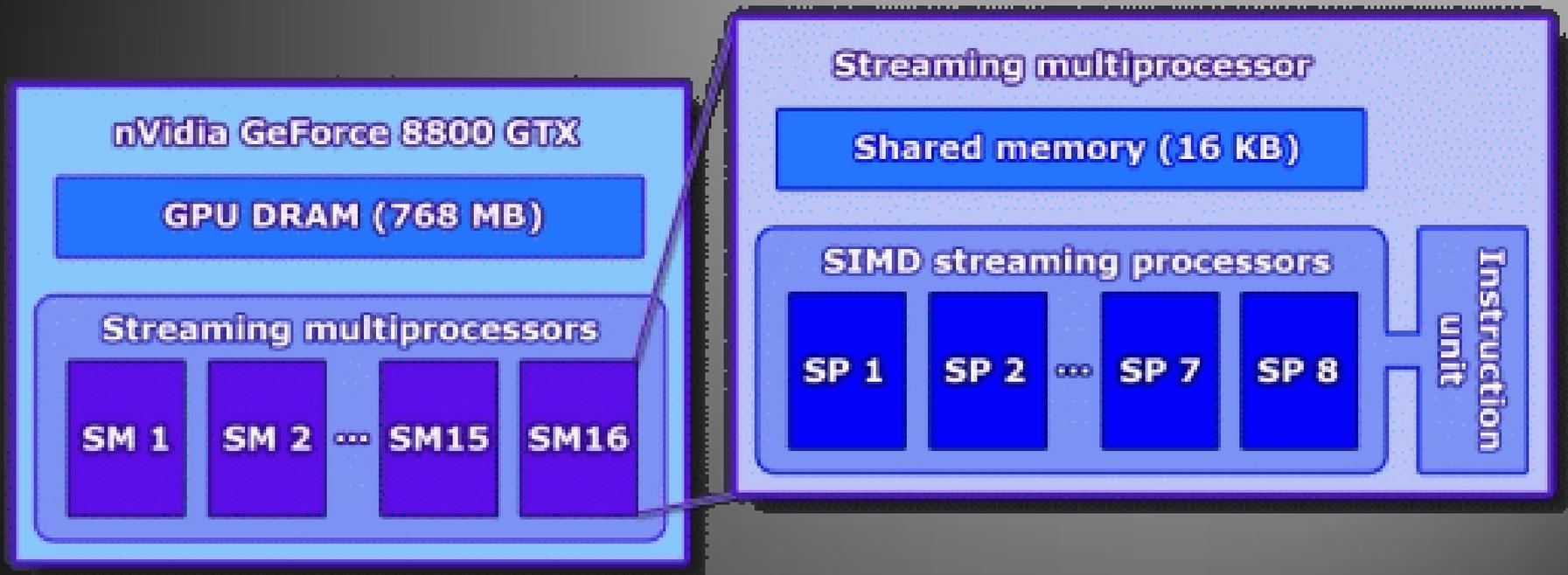
Conical Intersection & MECI



Relaxation & Barriere



Aufbau Grafikkarte



Two-electron repulsion integral

$$\chi_{\mu}(\vec{r}) \propto (x - x_{\mu})^l (y - y_{\mu})^m (z - z_{\mu})^n \exp\left(-\alpha_{\mu} |\vec{r} - \vec{R}_{\mu}|^2\right)$$

$$(\mu\nu | \lambda\sigma) = \iint \frac{\chi_{\mu}(\vec{r}_1) \chi_{\nu}(\vec{r}_1) \chi_{\lambda}(\vec{r}_2) \chi_{\sigma}(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2$$

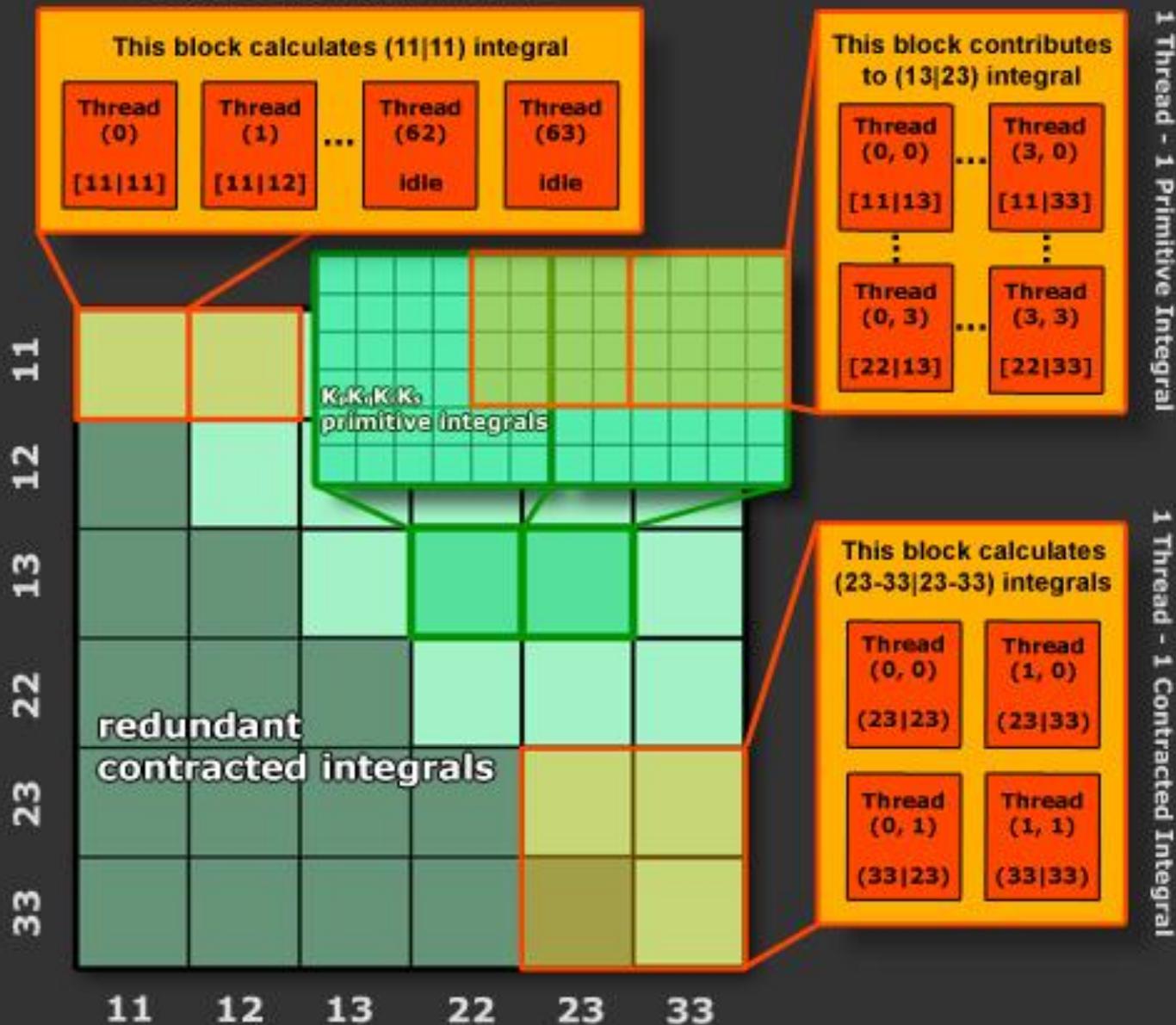
Orbitale bestehen aus Basisfunktionen, welche sich wiederum im Allgemeinen aus mehreren primitiven Funktionen von GTO's (Gaussian Type Orbital) zusammensetzen. Für eine Basisfunktion, die ein STO (Slater Type Orbital) darstellt, gilt daher:

$$(\mu\nu|\lambda\sigma) = \sum_{p=1}^{N_\mu} \sum_{q=1}^{N_\nu} \sum_{r=1}^{N_\lambda} \sum_{s=1}^{N_\sigma} d_{\mu p} d_{\nu q} d_{\lambda r} d_{\sigma s} [pq|rs]$$

Links steht ein Zwei-Elektronen-Integral über Basisfunktionen und rechts eine Summe über primitive Integrale, wobei die Koeffizienten vor dem Integral in eckigen Klammern konstant sind.

1T1CI (One Thread One Contracted Integral)	1B1CI (One Block one Contracted Integral)	1T1PI (One Thread one Primitive Integral)
Jeder Thread löst obige Gleichung.	Ein Block löst die obige Gleichung	Jeder Thread berechnet ein primitives Integral
→ Ergebnis wird in GPU DRAM gespeichert	Die 32 Threads erhalten zyklisch die zu berechnenden primitiven Integrale.	→ Ergebnisse werden in GPU DRAM gespeichert und darüber erfolgt die Addition der einzelnen primitiven Integrale, womit sich die Lösung der obigen Gleichung ergibt.
Problem: Unterschiedlich viele primitive Integrale, Summen also unterschiedlich lang und damit unterschiedlich lange Rechendauer	→ Aufsummation im gemeinsamen Speicher liefert Lösung der Gleichung, dieses wird im GPU DRAM gespeichert.	Problem: Zugriff auf GPU DRAM im Vergleich sehr langsam → Wenn obige Formel aus vielen Summanden besteht, deutliche Geschwindigkeitsnachteile
Verbesserung: Threads nach Anzahl der Primitiven Integrale sortieren, Programmieraufwand und Dauer der Umsortierung sind dabei relativ kurz	Problem: (Extremfall, wenn alle 4 Summen zusammenbrechen) Das contracted Integral besteht aus einem primitiven Integral → Restliche 31 Threads eines WARPS im SIMD Betrieb führen keine Aufgabe aus.	

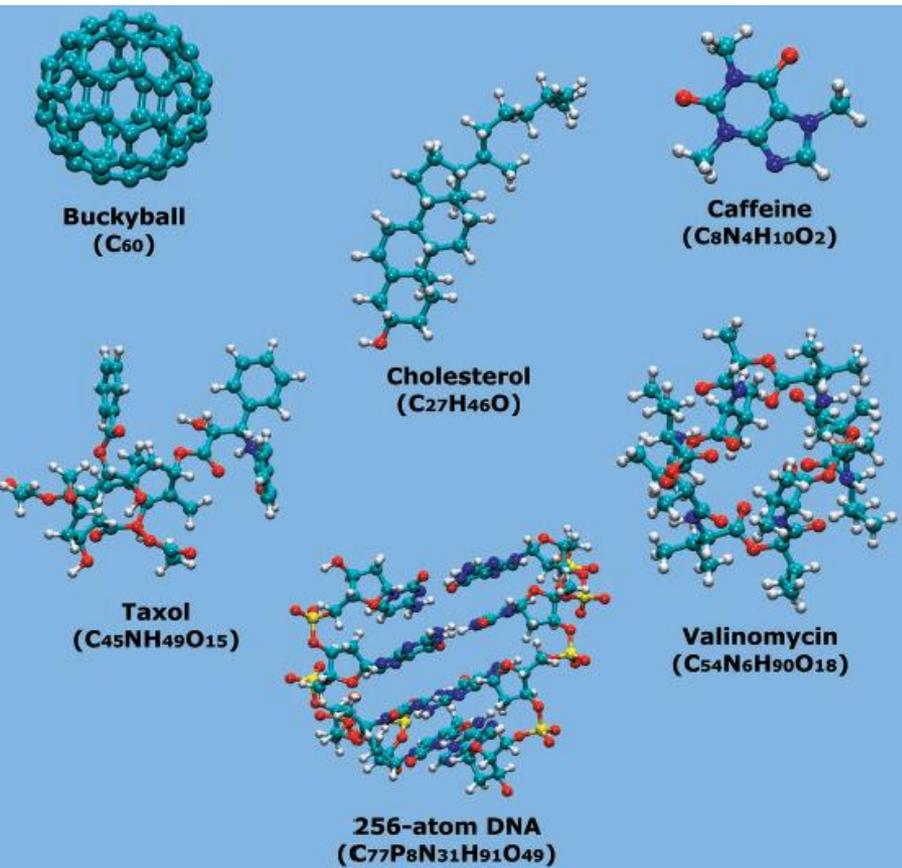
1 Block- 1 Contracted Integral



Vergleich der drei Methoden

Basis set	GPU (seconds)			CPU pre-calculation (seconds)	GPU-CPU transfer (seconds)	GAMESS # (seconds)
	1B1CI	1T1CI	1T1PI			
6-311G	7.086	0.675	0.428	0.009	0.883	170.8
STO-6G	1.608	1.099	2.863	0.012	0.012	90.8

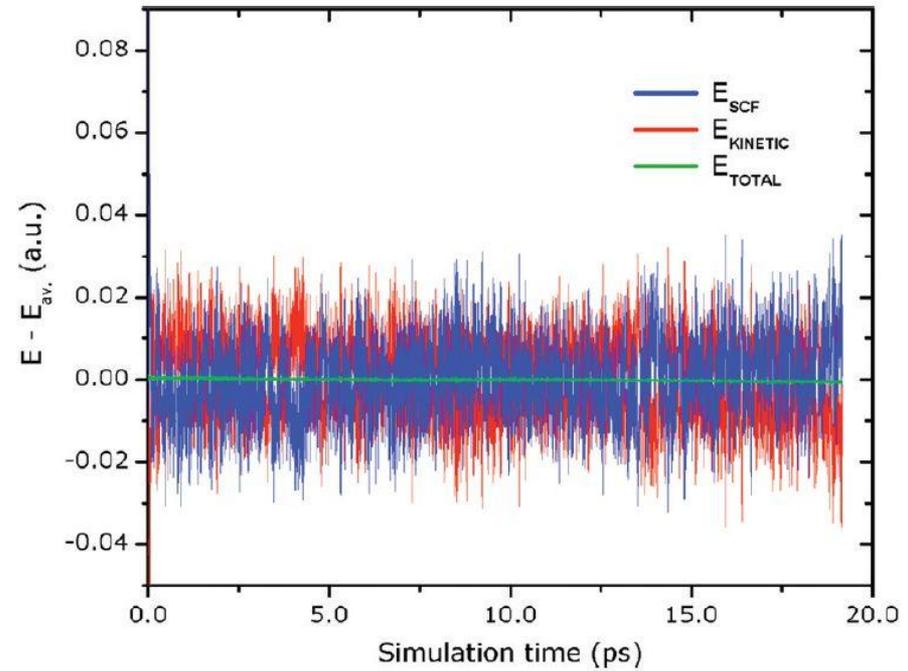
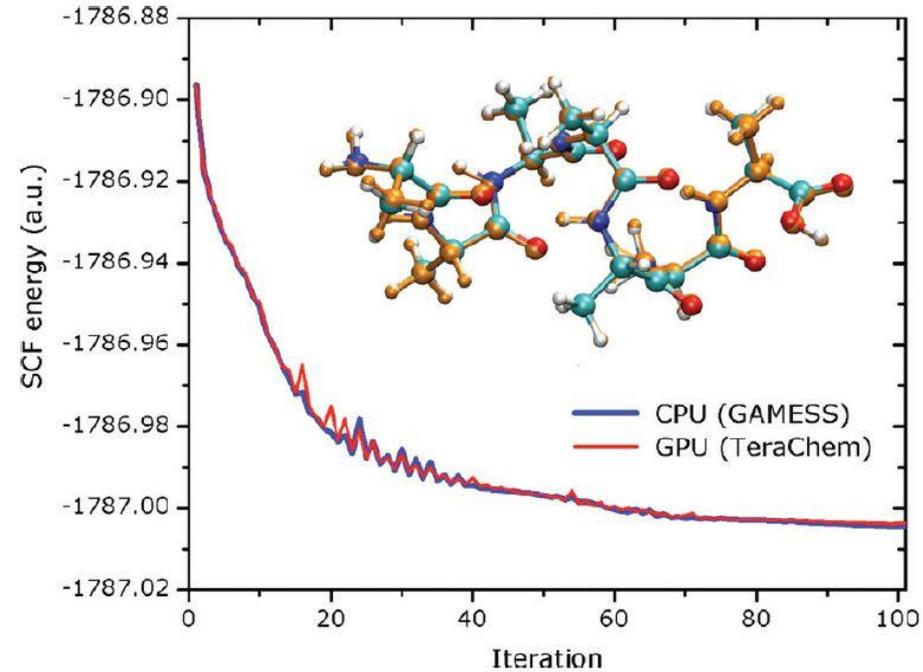
QC auf GPU



Molecule	Time per direct SCF iteration (seconds)	
	GPU	GAMESS
Caffeine	0.168	4.4
Cholesterol	1.23	68.0
Buckyball	5.71	332.0
Taxol	4.45	279.6
Valinomycin	8.09	750.6

Molecule	Electronic energy (atomic units)	
	GPU (32 bit)	GAMESS
Caffeine	-1605.91830	-1605.91825
Cholesterol	-3898.82158	-3898.82189
Buckyball	-10521.6414	-10521.6491
Taxol	-12560.6840	-12560.6828
Valinomycin	-20351.9855	-20351.9904

Konvergenz & Drift



Inhaltsverzeichnis

- ▶ **Einleitung**
 - Geschichte des Spiels
 - Entwicklung Videospiele
 - Bedeutung der Grafikkarte
- ▶ **Hauptteil**
 - Anwendungen in der Teilchenphysik
 - QC on GPU
- ▶ **Schluss**
 - Kommerzielle Verwendung in Forschung (QC)

Features

TeraChem is general purpose quantum chemistry software designed to run on NVIDIA GPU architectures under a 64-bit Linux operating system. Some of TeraChem features include:

- Full support for both NVIDIA Tesla and Fermi GPUs
- Restricted and Unrestricted Hartree-Fock and grid-based Kohn-Sham energy and gradient calculations
- Various DFT functionals (BLYP, B3LYP, PBE, revPBE) and DFT grids (800 - 80,000 grid points per atom)
 - Static and dynamical DFT grids
 - Empirical dispersion correction (DFT-D3 and DFT-D2)
- Geometry optimization and transition state search (including constraints)
- Ab initio molecular dynamics (NVE, NVT ensembles)
 - Reversible Born-Oppenheimer dynamics
 - Spherical boundary conditions
- Support of multiple-GPU systems
- Single/Dynamical/Double precision accuracy
- Up to 1000 times faster than conventional CPU-based codes
- Designed for large molecules - reads/writes PDB files directly
- QM/MM treatment of surrounding water molecules using TIP3P force field
- Natural bond orbital analysis through integration with NBO6

Kosten

TeraChem now shipping!

TeraChem v1.4	Academic	Industry / Govt
1-2 machines*	\$900/each	\$1800/each
3-4 machines*	\$630/each	\$1260/each
5-8 machines*	\$450/each	\$900/each
9-16 machines*	\$315/each	\$630/each

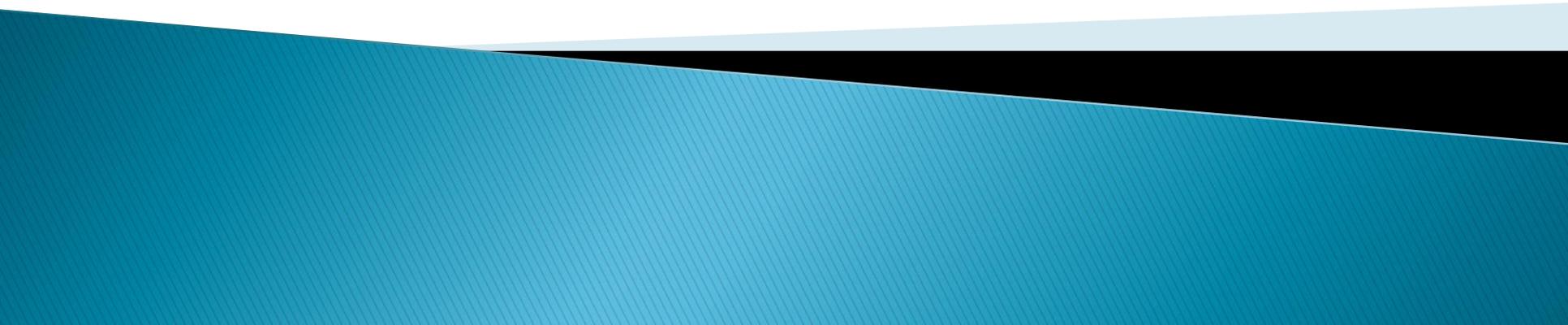
Tesla C1060 Computing Processor





<i>Processor</i>	1 x Tesla T10
<i>Number of cores</i>	240
<i>Core Clock</i>	1.33 GHz
<i>On-board memory</i>	4.0 GB
<i>Memory bandwidth</i>	102 GB/sec peak
<i>Memory I/O</i>	512-bit, 800MHz GDDR3
<i>Form factor</i>	Full ATX: 4.736" (H) x 10.5" (L) Dual slot wide
<i>System I/O</i>	PCIe x16 Gen2
<i>Typical power</i>	160 W

Primo Circenses, Inde Panis



Literatur

- ▶ http://de.wikipedia.org/wiki/%C3%9Cber_die_%C3%A4sthetische_Erziehung_des_Menschen
- ▶ http://de.wikipedia.org/wiki/Homo_ludens
- ▶ <http://de.wikipedia.org/wiki/Spieltheorie>
- ▶ http://de.wikipedia.org/wiki/Geschichte_der_Videospiele
- ▶ <http://de.wikipedia.org/wiki/GPGPU>
- ▶ „Grafikkarten für die Datenflut“ V. Lindenstruth. *Physik Journal* 10 (2011).
- ▶ „Photodynamics in Complex Environments: Ab Initio Multiple Spawning Quantum Mechanical/Molecular Mechanical Dynamics“ A.M. Virshup, C. Punwong, T.V. Pogorelov, B. Lindquist, C. Ko and T.J. Martínez, *J. Phys. Chem.*, Invited centennial feature article, 113B, 3280–3291 (2009).
- ▶ „Ab Initio Multiple Spawning Dynamics“ Todd Martinez. [\[pdf\]](#)
- ▶ „Introduction to Computational Chemistry“ F. Jensen. 2. Auflage. John Wiley & Sons 2007
- ▶ „Quantum Chemistry on Graphical Processing Units. 2. Direct Self-Consistent Field Implementation“ I.S. Ufimtsev and T.J. Martínez, *J. Chem. Theo. Comp.*, 5, 1004–1015 (2009).
- ▶ <http://www.petachem.com/>
- ▶ http://www.nvidia.de/object/computational_chemistry_de.html

GAME OVER

