Science & Technology Facilities Council

Technical Report
RAL-TR-2016-018

Proceedings of the MDANSE 2016 Workshop. Held at Cosener's House, Abingdon, UK, November 2016

S Mukhopadhyay, MA González, K Refson, J Armstrong, SF Parker, M Probert, N Hine, N Draper, E Pellegrini, G Kneller, MR Johnson, F Fernandez-Alonso

November 2016

©2016 Science and Technology Facilities Council

This work is licensed under a Creative Commons Attribution 3.0 Unported License.

Enquiries concerning this report should be addressed to:

RAL Library
STFC Rutherford Appleton Laboratory
Harwell Oxford
Didcot
OX11 0QX

Tel: +44(0)1235 445384
Fax: +44(0)1235 446403
email: libraryral@stfc.ac.uk

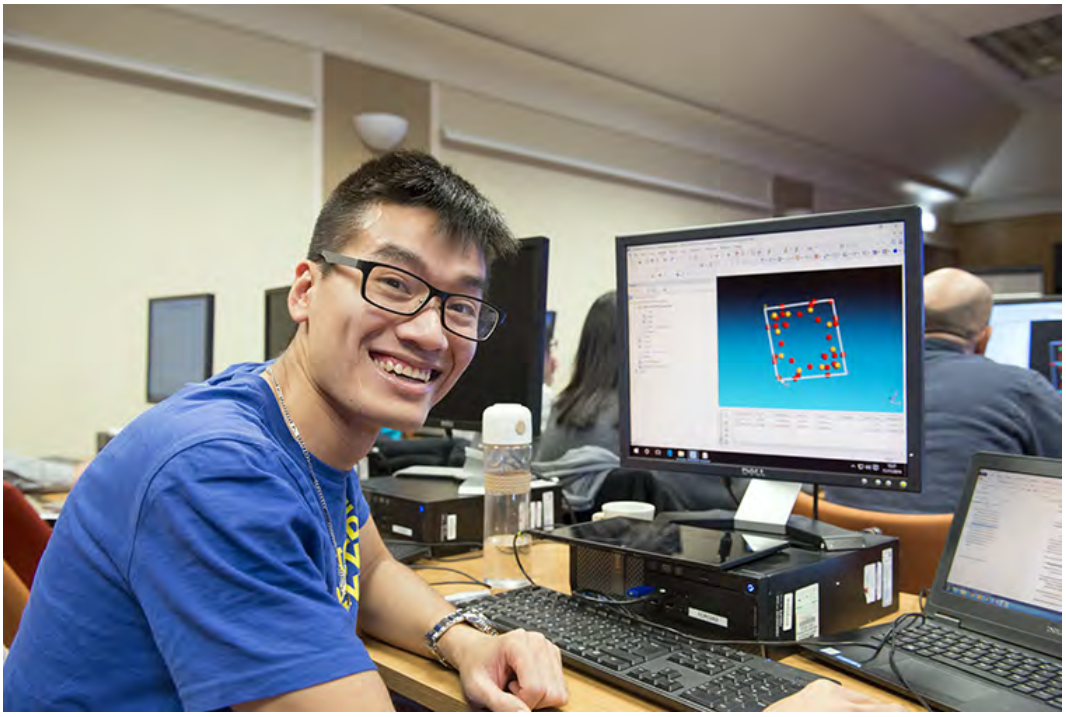
Science and Technology Facilities Council reports are available online at: http://epubs.stfc.ac.uk

ISSN 1358-6254

Neither the Council nor the Laboratory accept any responsibility for loss or damage arising from the use of information contained in any of their reports or in any communication about their tests or investigations.

Proceedings of the

MDANSE 2016 Workshop



Contributors:

S. Mukhopadhyay, M.A. González, K. Refson, J. Armstrong,
S.F. Parker, M. Probert, N. Hine, N. Draper, E. Pellegrini, G. Kneller
M.R. Johnson and F. Fernandez-Alonso

Foreword

In a colourful Autumn of November 2016, the *ISIS Molecular Spectroscopy Group* [1] proudly hosted, for the first time, *Molecular (and Lattice) Dynamics to Analyse Neutron Scattering Experiments – MDANSE 2016* at the Coseners House, Abingdon, United Kingdom. This workshop was immediately preceded by the *Molecular Spectroscopy Science Meeting – MSSM 2016* [2], and was jointly organised with the *ILL Computing for Science Group* [3], reflecting a long-lasting partnership between ILL and ISIS across the English Channel. The aim of this workshop was to show how widely available simulation tools could be applied to the analysis and interpretation of neutron-scattering data, as well as to bridge the knowledge gap between scientists coming from either experimental or computational backgrounds. With 50 delegates from around the globe, the event was a sell-out, attracting a balanced mix of attendees both from computational and experimental backgrounds. This event also provided a forum to discuss current and emerging challenges associated with the interpretation of neutron-scattering experiments and networking with experts in this field.

The programme of this three-day workshop focused on the two main computational methodologies currently in use to analyse neutron-spectroscopy experiments: Molecular Dynamics (MD) for the interpretation of QuasiElastic Neutron Scattering (QENS) spectroscopy, and Lattice Dynamics (LD) for Inelastic Neutron Scattering (INS). Both traditional force fields and state-of-the-art electronic-structure calculations, based on Density Functional Theory (DFT) were at the heart of this programme.

The workshop began with a warm welcome from Felix Fernandez-Alonso, Head of the *ISIS Molecular Spectroscopy Group*. This opening presentation emphasized the importance of using state-of-the-art computational methods to get the most out of experimental neutron data on increasingly complex materials and experimental conditions. These range from emerging high-throughput QENS and INS studies to subtle nuclear-quantum effects in condensed matter using neutron Compton scattering, all of which constitute an integral part of the current science & instrumentation programmes at ISIS [4, 5, 6] ISIS Catalysis Scientist, Stewart Parker gave an overview of the virtuous *structure-spectroscopy-theory* cycle and presented several examples of how DFT-based LD calculations could provide an in-depth understanding of experiments on catalysis and energy materials. After Stewart's talk Miguel González from the ILL made his presentation on fundamental theory of MD and force fields. He also compared capabilities and limitations of force-field-based and DFT-based methods.

The second set of talks in the afternoon of the first day focused on DFT and LD. In the first talk of this session, Sanghamitra Mukhopadhyay from ISIS presented the fundamental theory of electronic structure using DFT, with an emphasis on the need to move beyond their use as a 'black-box'. Keith Refson from ISIS followed, giving an overview of LD methods to calculate vibrational spectra for direct comparison to experiment. Keith's talk provided several detailed examples of applications of LD to interpret vibrational data across physical and chemical sciences, with an emphasis on INS.

The second day of the workshop was devoted entirely to hands-on computational tutorials supported by real experimental data, some of which are freely available from the *INS Database* [7]. Following a short introduction of the software tools by Miguel and Keith, the practical sessions were formally subdivided into the two main methodologies of MD and LD. Delegates were given the opportunity to choose the session which was most relevant to their research interests, and we were quite impressed to see that many of them managed to hop across MD and LD in order to gain an appreciation for both of them. Delegates were provided with step-by-step guidelines to complete the tutorials. They were also supplied with pre-computed simulations for post-processing and accompanying experimental data for comparison. All of these materials can be found online [8]. Delegates were encouraged to work with their own experimental data. These tutorials were highly interactive. Experts from both ISIS and ILL provided one-on-one tuition not only on the set tutorial problems, but also on the delegates' own research problems. Delegates used this opportunity to explore alternative methods that could be used. We held a session to explore a number of joint initiatives in the future. An enjoyable tour of the ISIS Facility in the early evening marked the end of the day.

The final day of the workshop was devoted to emerging developments and associated methodologies. Unfortunately, Gerald Kneller from CNRS-Orlèans could not make it in person to the event; his extensive presentation of novel ways to analyse MD simulations is available in this report. To make up for Gerald's absence Miguel stepped up to the podium and explained a different perspective to the analysis of MD simulations, including ways to validate and compare these with experimental results. He emphasized the capabilities of the MDANSE software [9] used in the tutorial sessions earlier. Matt Probert from the University of York gave an authoritative talk on Path-Integral Molecular Dynamics (PIMD), an essential tool to understand at a quantitative level the quantum nature of nuclear motions in condensed-matter systems. With a series of examples, Matt also explained how PIMD could be performed with the DFT code CASTEP [10], a software package that had been introduced in the tutorial session of the previous day. Nick Hine of the University of Warwick continued the session with a detailed presentation of advanced simulation methods based on linear-scaling DFT. Nick emphasized the importance of the method to simulate large systems up to tens of thousands of atoms and involving both order and disorder at the nanoscale. Finally, Nick Draper from ISIS and Tessella introduced the MANTID framework [11], a comprehensive and novel software package focused on the reduction, visualisation, and analysis of neutron and muon data.

As MDANSE2016 came to a close, Felix Fernandez-Alonso provided a few final thoughts, with an emphasis on developing increasingly stronger platforms and initiatives for discussion, exchange, and joint efforts across experiments and simulations in the context of neutron scattering.

The workshop was not all work only, delegates did walk out of the venue to enjoy an informal dinner in a stylish local pub at the end of the first day. A more formal three course dinner was served to the delegates at the Cosener's house on

the second day. From their relaxed mood there was no doubt that they very much enjoyed it all.

Finally, we would like to take this opportunity to thank Emma Roberts of ISIS Users office for helping in logistical arrangements of the workshop. Without her professional and sincere help this workshop wouldn't have been such a success. We also would like to sincerely thank Tom Griffin and the members of his ISIS Computing group for transforming the Garden room at the Cosener's house to a vibrant computing lab, the ISIS Communication group for helping in arranging the ISIS tour and the colleagues of the Molecular spectroscopy group for helping in the demonstration of the ISIS experimental hall. We owe a great deal to Stephen Kill for all the nice photos that will help keep the good memories on record, and the staff of the Cosenar's House for their help and co-operation. We look forward to seeing you again in a future MDANSE workshop!

The Organising Committee
November 2016

References

[1] *www.isis.stfc.ac.uk/groups/molecular-spectroscopy*

[2] See *www.isis.stfc.ac.uk/news-and-events/events/2016*. The proceedings of this meeting will also be available from *epubs.stfc.ac.uk* in due course.

[3] *www.ill.eu/instruments-support/computing-for-science/home*

[4] S. Rudić, A.J. Ramirez-Cuesta, S.F. Parker, F. Fernandez-Alonso, R.S. Pinna, G. Gorini, C.G. Salzmann, S.E. McLain, and N.T. Skipper, TOSCA International Beamline Review, *Rutherford Appleton Laboratory Technical Report RAL-TR-2013-015* (Didcot, 2013). Available online at *epubs.stfc.ac.uk/work/11216706*

[5] F. Demmel, V. Garcia-Sakai, S. Mukhopadhyay, S.F. Parker, F. Fernandez-Alonso, J. Armstrong, F. Bresme, J.P. de Vicente, F. Sordo, F.J. Bermejo, C.G. Salzmann, and S.E. McLain, IRIS & OSIRIS International Beamline Review – Report Presented to the Panel, *Rutherford Appleton Laboratory Technical Report RAL-TR-2014-014* (Didcot, 2014). Available online at *epubs.stfc.ac.uk/work/12271760*

[6] A.G. Seel and R. Senesi eds, *J. Phys. Conf. Ser.* **571** (2014). This special issue is available online at *iopscience.iop.org/1742-6596/571/1*

[7] *www.isis.stfc.ac.uk/instruments/tosca/ins-database*

[8] *DOI:http://doi.org/10.5281/zenodo.204818*

[9] MDANSE can be downloaded free of charge from *mdanse.org/*

[10] *www.castep.org/*

[11] MANTID can be downloaded free of charge from *www.mantidproject.org*

Molecular (and Lattice) Dynamics to Analyse Neutron Scattering Experiments

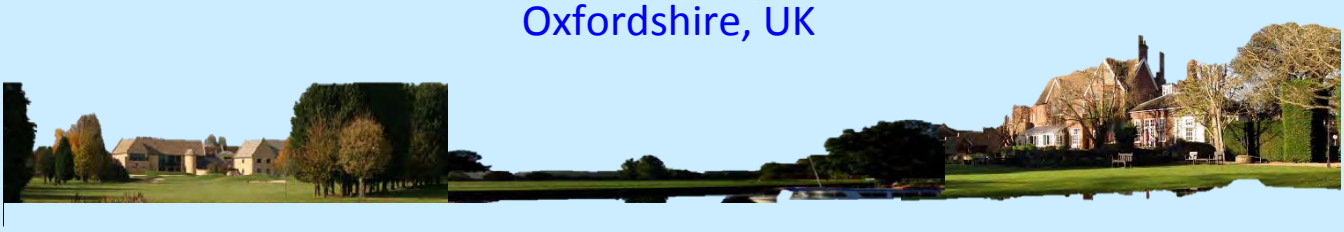
MDANSE 2016

10th -12th November

Science & Technology Facilities Council
ISIS

The Cosener's House
Abingdon
Oxfordshire, UK

NEUTRONS FOR SCIENCE



The *ISIS Molecular Spectroscopy Group* and *ILL Computing for Science Group* will hold the next MDANSE (Molecular Dynamics and Lattice Dynamics to Analyse Neutron Scattering Experiments) workshop at the Cosener's House, Abingdon, UK from the 10th to the 12th of November 2016.

The goal of this school is to show how widely-available simulation tools can be applied to the analysis and interpretation of neutron scattering data. The school will last for three days and will concern mainly the practical aspects of running and analysing simulations so that, numerical and experimental data can be directly compared and simulations can then be investigated in atomic detail.

The focus will be on the use of Molecular Dynamics and lattice dynamics simulations to explore diffusive processes on the time scales covered by low energy neutron scattering spectrometers, phonon dispersion curves and vibrational densities of states that are typically measured with low energy and inelastic neutron scattering spectrometers.

The school is designed for providing practical training, so the number of participants is limited to 40. In order to select candidates, we ask applicants to state briefly their scientific fields of interest, why they would like to attend the school and previous experience with computer simulations.

This year's workshop will focus on the following areas

Molecular Dynamics

Lattice Dynamics

First-Principles Simulations

Hands on Tutorials

Invited speakers include

- G. Kneller, CNRS, France
- Matt Probert, University of York, UK
- Nicholas Hine, University of Warwick, UK

Fees: A single fee of £100 is requested.

(This fee will cover the accommodation and meal expenses)

Registration is now open and will close on 31st August.

To register, please click here

We look forward to seeing you in November!

Organisers:
Sanghamitra Mukhopadhyay, ISIS
Miguel Gonzalez, ILL
Jeff Armstrong, ISIS
Keith Refson, ISIS
Mark Johnson, ILL
Felix Fernandez-Alonso, ISIS

MDANSE 2016 Workshop

Time	Thursday, 10th November	Friday, 11th November	Saturday, 12th November
9.00		**Introduction to software to be used in tutorials** Miguel Gonzalez, Keith Refson	**Advanced Analysis in Computer Simulations (nMoldyn)** Gerald Kneller
9.30		**Hands on tutorials** Two groups will work simultaneously 1.Molecular Dynamics and QENS 2. Lattice Dynamics and INS Jeff/Miguel/Sanghamitra/Keith	
9.45			**Path Integral MD (PIMD)** Matt Probert
10.00			
10.30			**Refreshments**
11.00	Registrations	**Refreshments**	**Linear scaling DFT and Applications (OneTep)** Nicholas Hine
		Hands on tutorials (Continued)	
11.45			**Mantid for Data Analysis** Nick Draper
12.00	**Lunch**		
12.30		**Group Photo**	**Wrapping up** Felix Fernandez-Alonso
13.00		**Lunch**	**Lunch**
13.45	**Welcome** Felix Fernandez-Alonso		
14.00	**Neutron Scattering & Computer Simulations** Stewart Parker	**Hands on tutorials (Continued)**	
14.45	**Basic MD and Force fields** Miguel Gonzalez		
15.00			
15.30	**Refreshments**	**Refreshments**	
16.00	**Introduction to Electronic Structure** Sanghamitra Mukhopadhyay	**Hands on tutorials (Continued)**	
16.30	**Introduction to Lattice Dynamics** Keith Refson	**Round Table Discussions**	
17.00		**Making Way to Bus**	
17.15		**ISIS Tour**	
17.30	**Setting up Computers for the next day**		
18.00	**Break**		
19.00	**Dinner** **(Crown & Thistle)**		
		Workshop Dinner **(Cosener's House)**	
20.00			

List of registered participants

Name	Institution	Email	Expertise
Alex Aziz	University of Reading, United Kingdom	a.g.aziz@pgr.reading.ac.uk	MD, DFT,modelling thermal conductivity
Alexander Hawkins	University of Glasgow, United Kingdom	a.hawkins.1@research.gla.ac.uk	MD, QENS
Alexander O'Malley	Cardiff University & UK Catalysis Hub, United Kingdom	alexomalley@btinternet.com	MD, QENS, diffusion in porous material, catalysis
Anders Markvardsen	ISIS Pulsed Neutron & Muon Source, United Kingdom	anders.markvardsen@stfc.ac.uk	Mantid
Andrea Zachariou	University of Glasgow United Kingdom	ac.zachariou@gmail.com	MD, INS, QENS
Antony Phillips	Queen Mary University of London United Kingdom	a.e.phillips@qmul.ac.uk	Neutron diffraction, INS, QENS, Lattice dynamics
Asel Sartbaeva	University of Bath United Kingdom	a.sartbaeva@bath.ac.uk	INS, diffraction
Chiara Ferrara	Universit di Pavia, Itally	chiara.ferrara01@ateneopv.it	DLPOLY, GULP, neutron diffraction, SOFC, Li ion battery
Chris Ablitt	Imperial College London, United Kingdom	christopher.ablitt14@imperial.ac.uk	DFT
Chris Sewell	Imperial College London, United Kingdom	c.sewell14@imperial.ac.uk	DFT
Daniel Dervin	Queens University Belfast, United Kingdom	ddervin01@qub.ac.uk	MD, QENS
Dirk Visser	Louhborough University, United Kingdom	d.visser@lboro.ac.uk	Diffraction, spectroscopy
Dominik Zeller	Laboratoire Interdisciplinaire de Physique, Universit Grenoble Alpes, France	zeller@ill.fr	MD, INS, biomolecules
Felix Fernandez-Alonso	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	felix.fernandez-alonso@stfc.ac.uk	physical & materials chemistry, spectroscopy, neutron scattering, computational modelling
Frederic Ngono	ILL (Grenoble) / Universite de Lille 1 France	ngonof@ill.fr	MD, Dmol3, neutron diffraction and spectroscopy, amorphous system
Gerald Kneller	CNRS - University of Orleans, France	gerald.kneller@cnrs.fr	nMoldyn, MD
Goran Svaland	Imperial College London, United Kingdom	g.svaland15@imperial.ac.uk	MD

List of registered participants (Cont.)

Name	Institution	Email	Expertise
Huan Doan	University of Bath, United Kingdom	dvh21@bath.ac.uk	MD, crystalographiy
Ian Silverwood	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	ian.silverwood@stfc.ac.uk	INS, QENS, calalysis
Jack Parsons	University of Southampton United Kingdom	jdparsons@hotmail.co.uk	INS, QENS, catalysis
Jeff Armstrong	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	jeff.armstrong@stfcl.ac.uk,	MD, LAAMPS, INS
Joaquin Klug	Queen's University Belfast United Kingdom	j.klug@qub.ac.uk	MD
Jonathan Fernando Gebbia	UniversitaPolitcnica de Catalunya Spain	jonathan.fernando.gebbia@upc.edu	DFT, INS
Juan David Olarte Plata	Imperial College London United Kingdom	j.olarte@imperial.ac.uk	MD, stochastic rotational dynamics
Kacper Drubicki	Adam Mickiewicz University &, JINR, Dubna , Poland & Russia	kacdru@amu.edu.pl	DFT
Keith Refson	Royal Holloway, University of London & ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	Keith.Refson@rhul.ac.uk	DFT, lattice dynamics
Krzysztof Dymkowski	Scientific Computing Department Rutherford Appleton Laboratory, United Kingdom	krzysztof.dymkowski@stfc.ac.uk	DFT, software development, Mantid
Louise McCann	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	louise.mccann@stfc.ac.uk	Software development, Mantid
Maria Silvina Moyano	Queen's University Belfast United Kingdom	m.moyano@qub.ac.uk	MD, biomolecules
Matt Probert	University of York United Kingdom	matt.probert@york.ac.uk	PIMD, DFT, ab-initio MD
Matthew Krzystyniak	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	matthew.krzystyniak@stfc.ac.uk	Neutron Compton scattering, NMR, DFT
Matthew Ryder	University of Oxford, United Kingdom	matthew.ryder@eng.ox.ac.uk	DFT, INS, MOF
Matthias Gutmann	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	matthias.gutmann@stfc.ac.uk	Neutron and X-ray crystallography, diffuse scattering, disordered materials
Michael Ferguson	Queen's University Belfast United Kingdom	mferguson19@qub.ac.uk	MD, DFT

List of registered participants (Cont.)

Name	Institution	Email	Expertise
Miguel A. Gonzalez	Institut Laue-Langevin France	gonzalezm@ill.fr	MD, MDANSE
Nadir Basma	Imperial College London & University College London, United Kingdom	n.basma@imperial.ac.uk	Neutron Diffraction, reverse Monte Carlo, solvent structures
Nick Hine	University of Warwick United Kingdom	n.d.m.hine@warwick.ac.uk	Linear scaling DFT Onetep
Nicholas Draper	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	nick.draper@stfc.ac.uk	Mantid
Ryan Kavanagh	Queens University Belfast United Kingdom	rkavanagh04@qub.ac.uk	
Sakiru Badmos	University College London United Kingdom	sakiru.badmos.14@ucl.ac.uk	MD
Sanghamitra Mukhopadhyay	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	sanghamitra.mukhopadhyay@stfc.ac.uk	DFT, MD, INS, QENS
Stephanie Chapman	University of Southampton United Kingdom	stephanie.chapman@soton.ac.uk	INS, QENS, catalysis zeolite and zeotype
Stewart Parker	ISIS Pulsed Neutron & Muon Source, Rutherford Appleton Laboratory, United Kingdom	stewart.parker@stfc.ac.uk	DFT, INS, catalysis
Tetiana Lemishko	Institut Laue-Langevin (ILL) France	lemishkot@ill.fr	DFT
Thomas Lorne	Institut Laue-Langevin (ILL) France	lorne@ill.fr	DFT, MD, MDANSE, INS
Tim Tejsner	Institut Laue-Langevin (ILL) France	ttejsner@gmail.com	
Tran Le	University College London United Kingdom	thi.le.15@ucl.ac.uk	MD, LAAMPS, GROMACS
Xiaoyu Han	University College London United Kingdom	x.han@ucl.ac.uk	DFT
Zheng Jiang	University of Southampton United Kingdom	z.jiang@soton.ac.uk	INS, catalysis

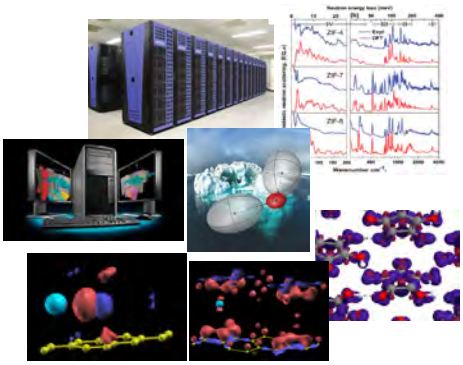
Then and Now

1990s

The 'Ball-and-stick' Way

21st Century

In-silico Neutron Scattering



Software packages have reached a level of sophistication whereby experimentalists can also enter the game.

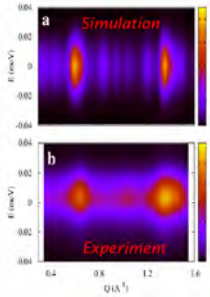
15 years ago, this task was much, much harder.

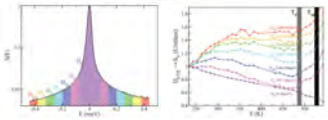
MDANSE School 2016

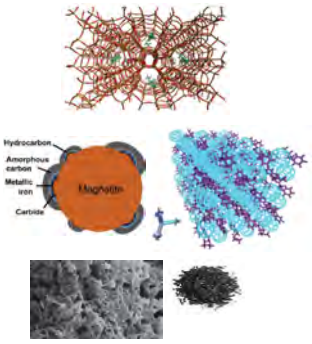
Time	Thursday, 10th November	Friday, 11th November	Saturday, 12th November
9.00		Introduction to software to be used in tutorials Miguel Gonzalez, Keith Refson	Advanced Analysis in Computer Simulations (nMoldyn) Gerald Kneller
9.30		Hands on tutorials Two groups will work simultaneously 1.Molecular Dynamics and QENS 2. Lattice Dynamics and INS Jeff/Miguel/Sanghamitra/Keith	
9.45			Path Integral MD (PIMD) Matt Probert
10.00			
10.30			Refreshments
11.00	Registrations	Refreshments	Linear scaling DFT and Applications (OneTep) Nicholas Hine
		Hands on tutorials (Continued)	
11.45			Mantid for Data Analysis Nick Draper
12.00	Lunch		
12.30		Group Photo	Wrapping up Felix Fernandez-Alonso
13.00		Lunch	Lunch
13.45	Welcome Felix Fernandez-Alonso		
14.00	Neutron Scattering & Computer Simulations Stewart Parker	Hands on tutorials (Continued)	
14.45	Basic MD and Force fields Miguel Gonzalez		
15.00			
15.30	Refreshments	Refreshments	
16.00	Introduction to Electronic Structure Sanghamitra Mukhopadhyay	Hands on tutorials (Continued)	
16.30	Introduction to Lattice Dynamics Keith Refson	Round Table Discussions	
17.00		Making Way to Bus	
17.15		ISIS Tour	
17.30	Setting up Computers for the next day		
18.00	Break		
19.00	Dinner (Crown & Thistle)	Workshop Dinner (Cosener's House)	
20.00			

In-silico Developments on the Horizon

HT QENS

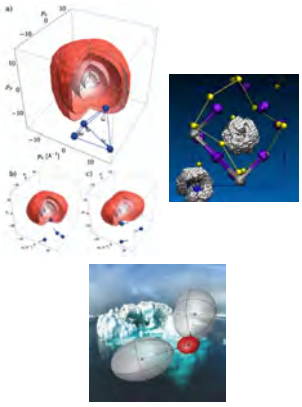






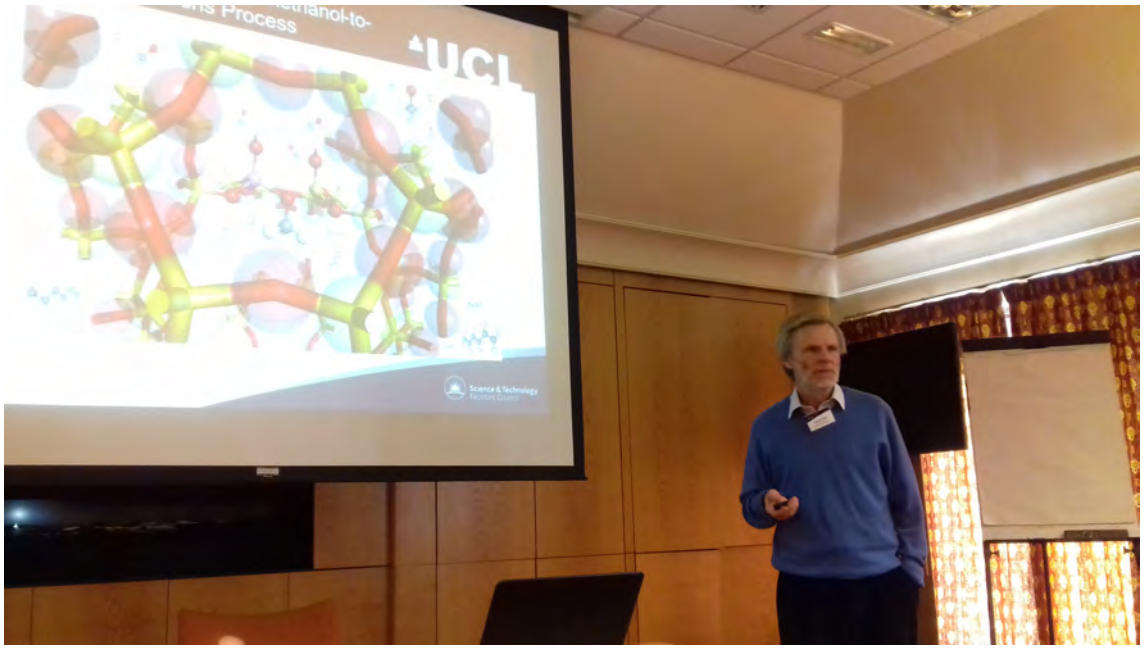
Disordered & Nanostructured Matter

Nuclear Quantum Dynamics



Key to move forward:

- *Your input & engagement.*
- *Resource.*

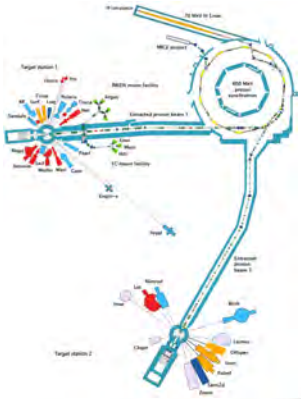

Process
UCL

Introduction to Electronic Structures
Sanghamitra Mukhopadhyay

in accuracy vs efficiency

Neutron scattering and Computer Simulations

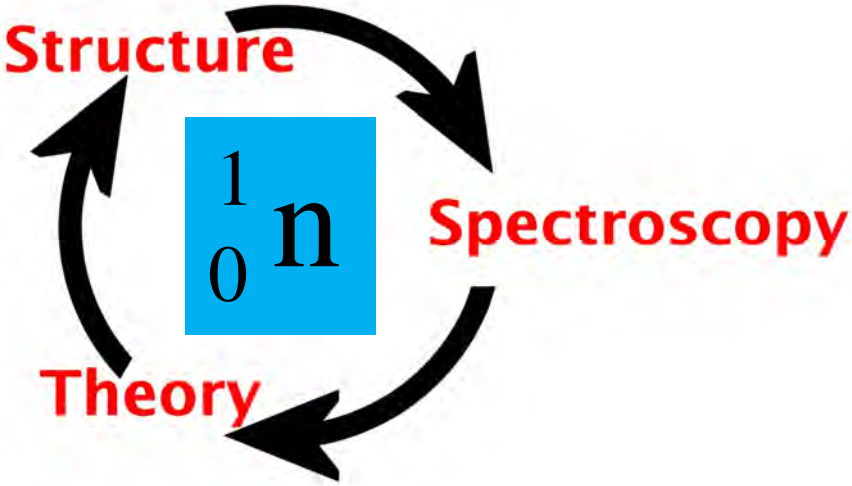
Stewart F. Parker
ISIS Facility



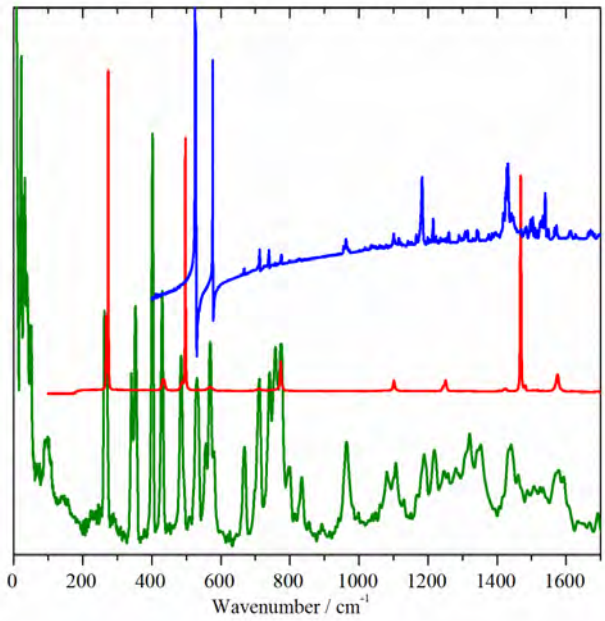
MDANSE
November 10th 2016
The Coseners House

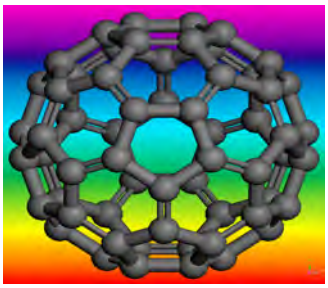
Science & Technology Facilities Council

ISIS



Science & Technology Facilities Council



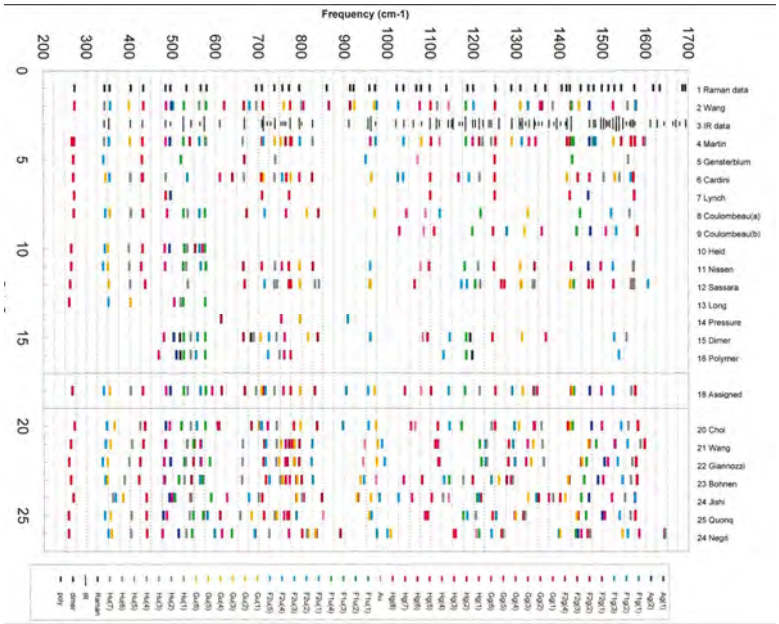


C_{60}

"The most beautiful molecule" (PCBM/P3HT)

Phys. Chem. Chem. Phys., 13 (2011) 7789 – 7804

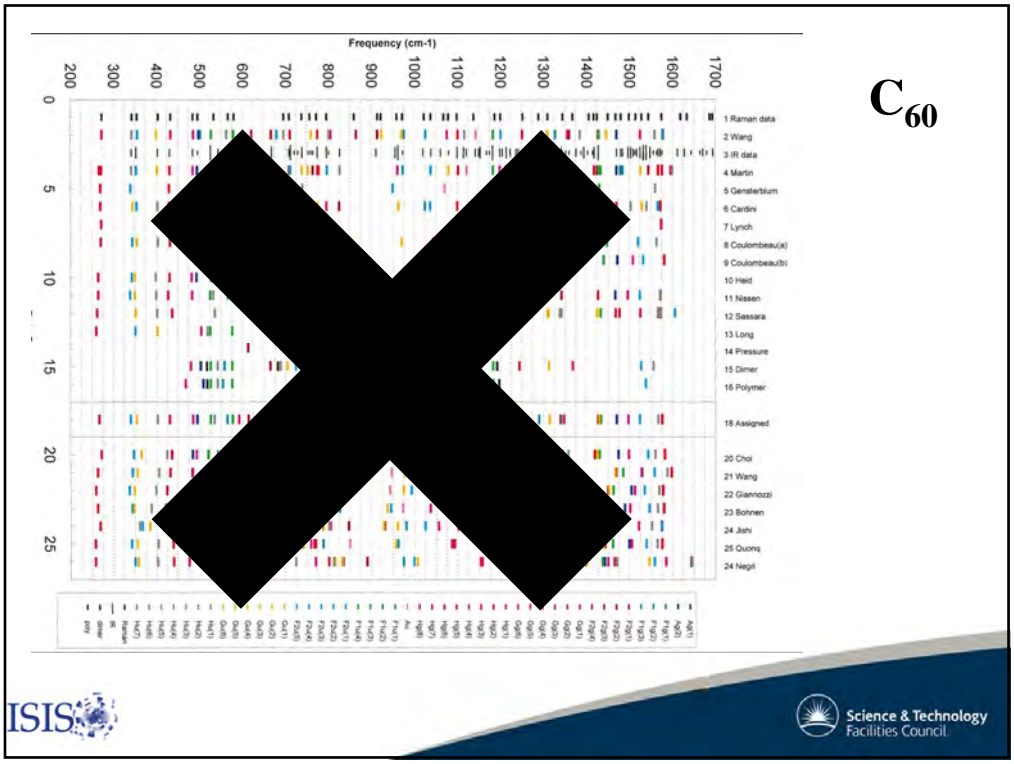
Science & Technology Facilities Council



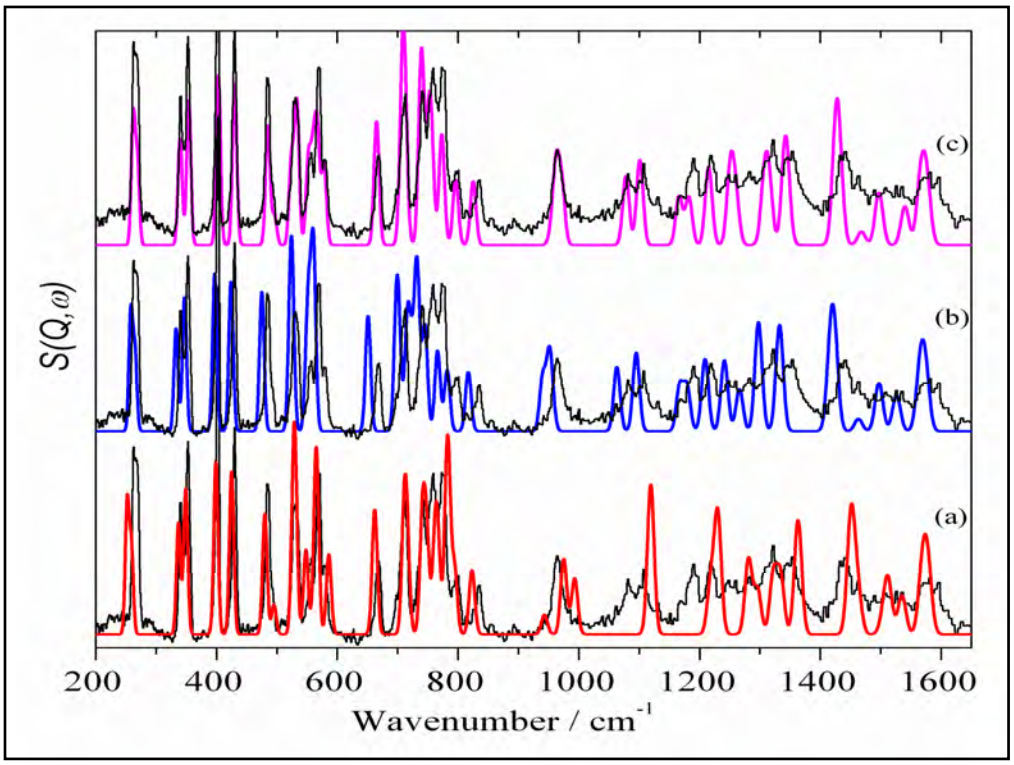
C_{60}

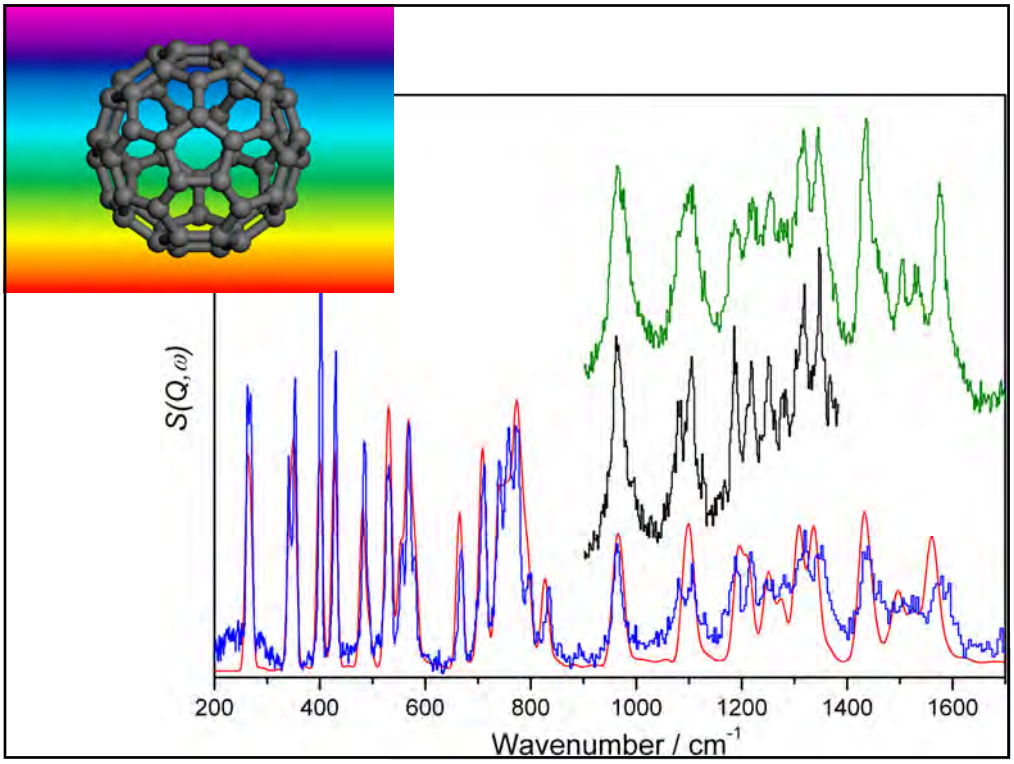
ISIS

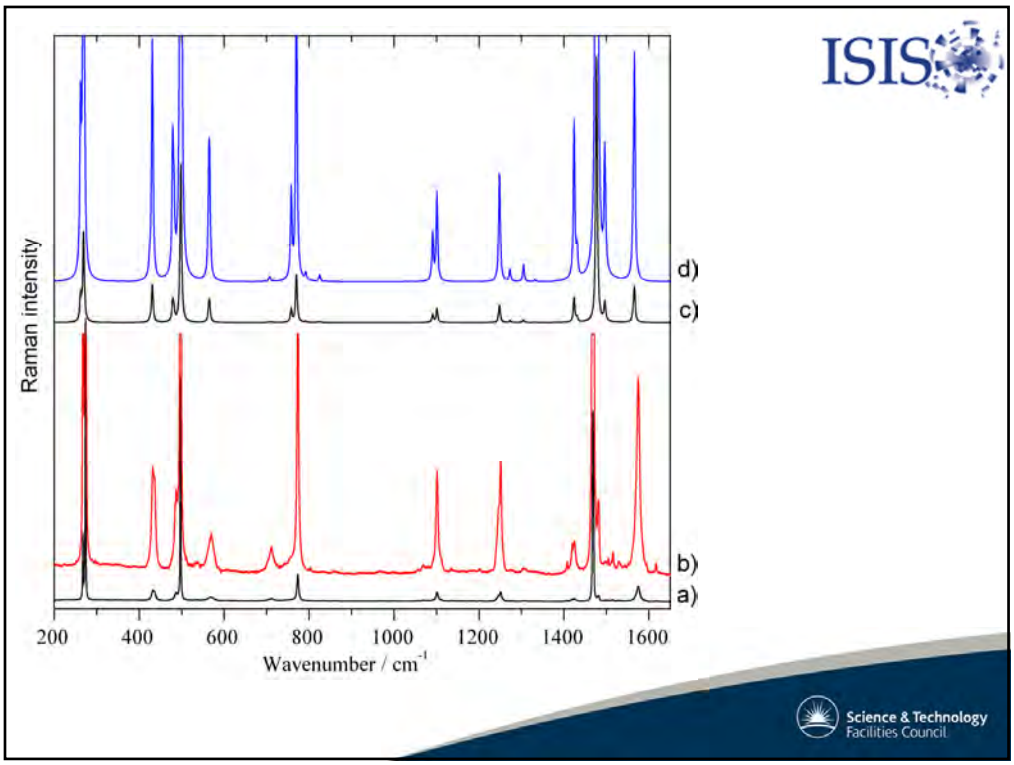
Science & Technology Facilities Council


Frequency (cm-1)
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
0
5
10
15
20
25
1 Raman data
2 Wang
3 IR data
4 Martin
5 Gensterblum
6 Cardini
7 Lynch
8 Coulombeau(a)
9 Coulombeau(b)
10 Heid
11 Nissen
12 Sassara
13 Long
14 Pressure
15 Dimer
16 Polymer
18 Assigned
20 Choi
21 Wang
22 Giannozzi
23 Bohnen
24 Jishi
25 Quong
24 Negri
ISIS
Science & Technology Facilities Council

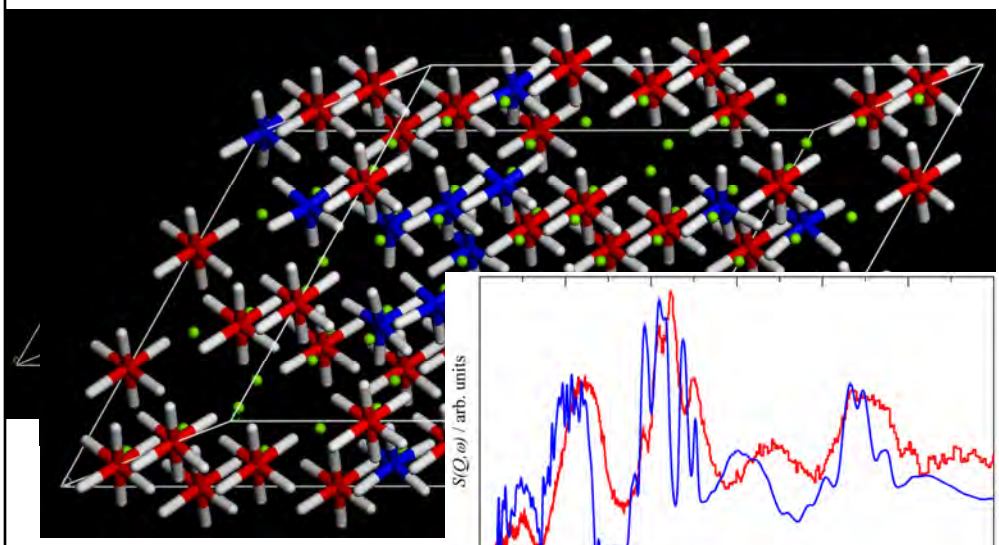
C_{60}


S(Q,ω)
(c)
(b)
(a)
200
400
600
800
1000
1200
1400
1600
Wavenumber / cm-1


S(Q, ω)
200
400
600
800
1000
1200
1400
1600
Wavenumber / cm-1

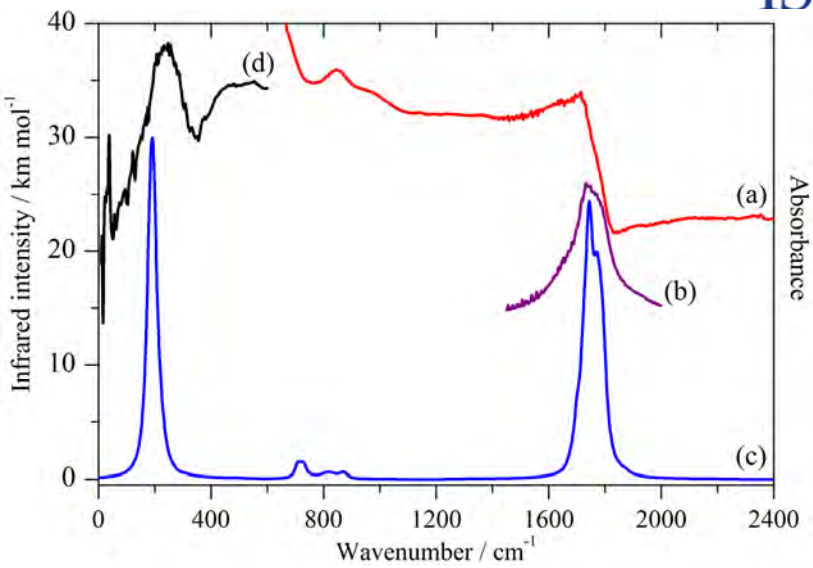

ISIS
Raman intensity
d)
c)
b)
a)
200
400
600
800
1000
1200
1400
1600
Wavenumber / cm-1
Science & Technology Facilities Council

$Mg_2[CoH_5]_{0.5}[FeH_6]_{0.5}$



34 ions 68 ions 306 ions

***Amm*2 2x2x2 4x3x3**



CuH: a (nearly) 200 year old problem

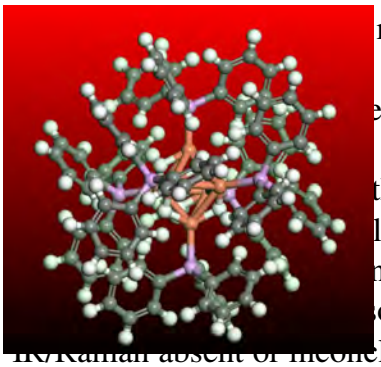
Charles-Adolphe Würtz
1817-1884

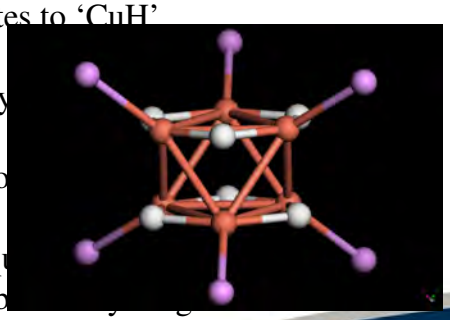
Science & Technology Facilities Council

Why CuH?

Only binary metal hydride made by solution chemistry
Intermediate phase in the dissolution of brass in sulphuric acid
Hydrogen is stored subsurface in $Cu/Zn/Al_2O_3$ methanol synthesis catalysts as CuH?
Mild reducing agent in organic syntheses, superseded by Stryker's reagent $[Ph_3PCuH]_6$

***i.e.* No consistent story**





routes to 'CuH'

ed by

th co

l

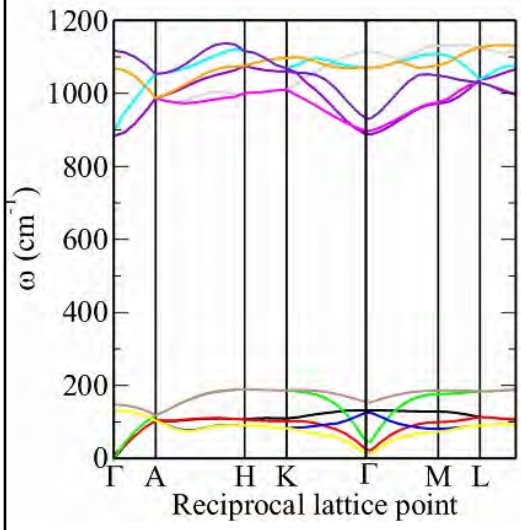
n-aqu

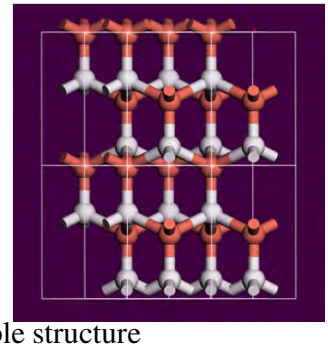
olub

IR/Raman absent or inconclusive

Science & Technology Facilities Council

Theory - CASTEP



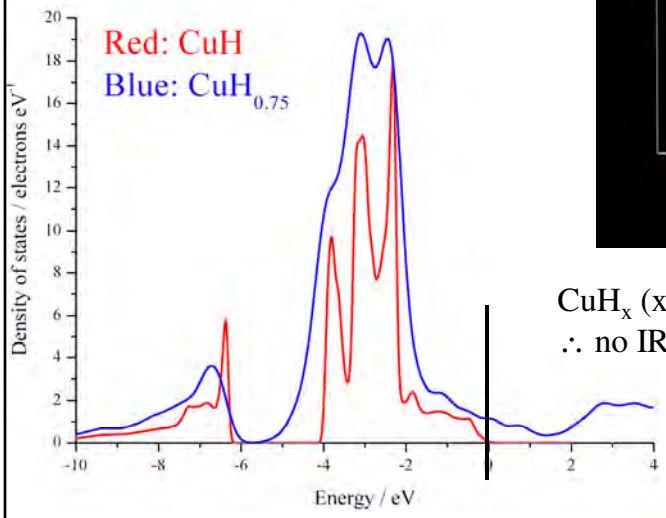


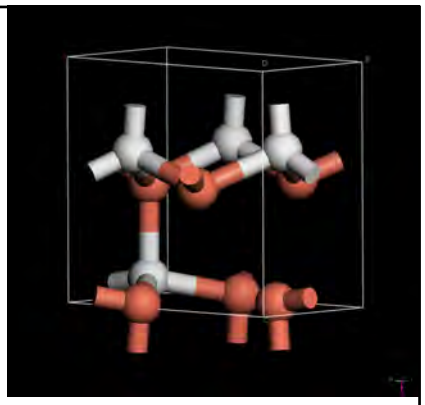
Stable structure
All positive frequencies
CuH is an insulator
Ionic solid: $Cu^{+0.37}$ $H^{-0.37}$
IR modes at: 883, 894 cm^{-1}
Raman modes at 883, 894,1068 cm^{-1}
Cu–H = 1.77 Å

Bennett *et al*,
***Inorg. Chem*. 54 (2015) 2213**
***Acta Cryst*. B71 (2015)**

Science & Technology Facilities Council

But: CuH is always non-stoichiometric,
$CuH_{0.6}$ to $CuH_{0.9}$
Model as $CuH_{0.75}$

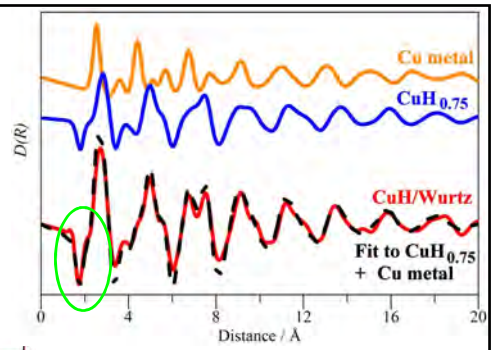


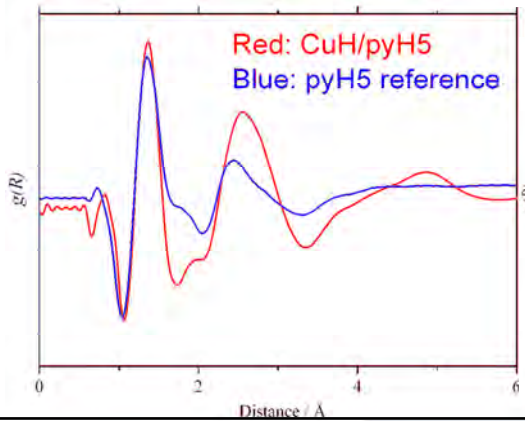


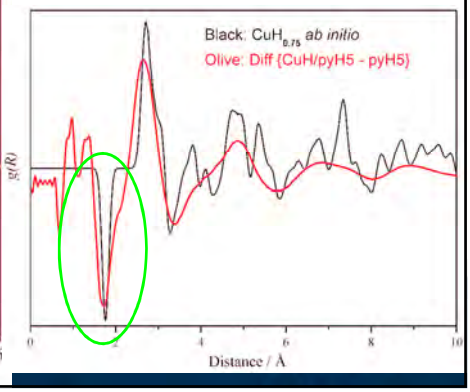
CuH_x (x<1) is a metal
∴ no IR or Raman active modes

Science & Technology Facilities Council

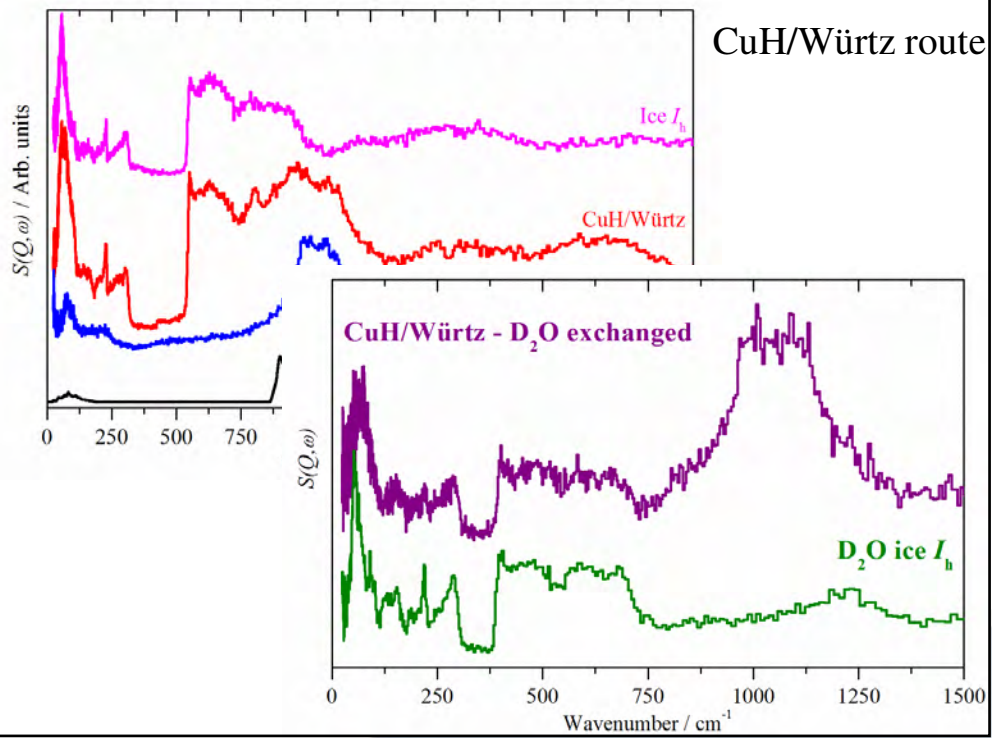
Question 1:
Do aqueous and non-aqueous routes produce the same CuH?

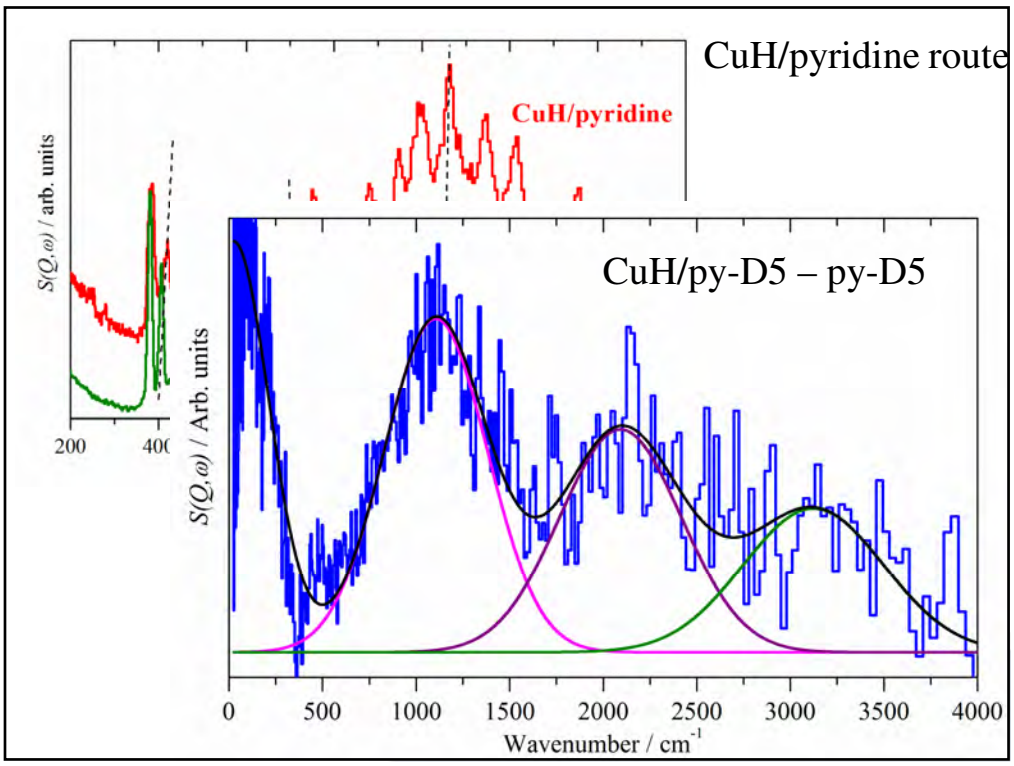






CuH/Würtz route





ISIS

Question 1: Do both the aqueous and non-aqueous routes produce the same CuH?

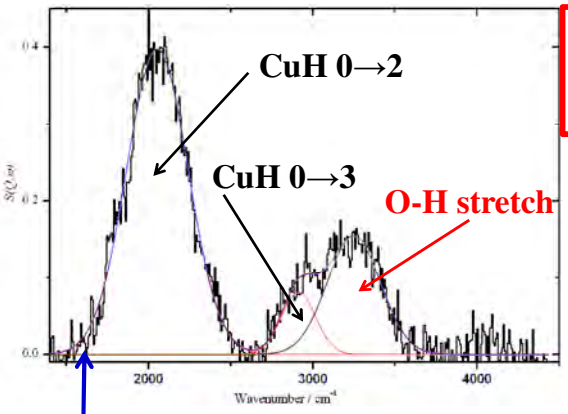
YES

The ready exchange of H_2O for D_2O and pyridine for pyridine-D5 while leaving the CuH unchanged validates the core-shell model.

The particle size is much smaller for the pyridine route than the other methods.

Science & Technology Facilities Council

Dried CuH/Würtz in the high energy region (MAPS)

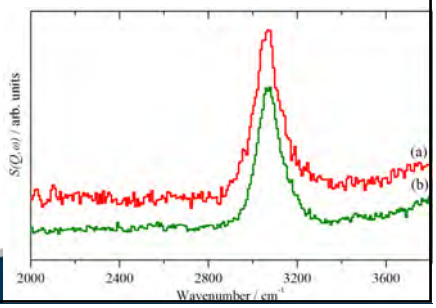


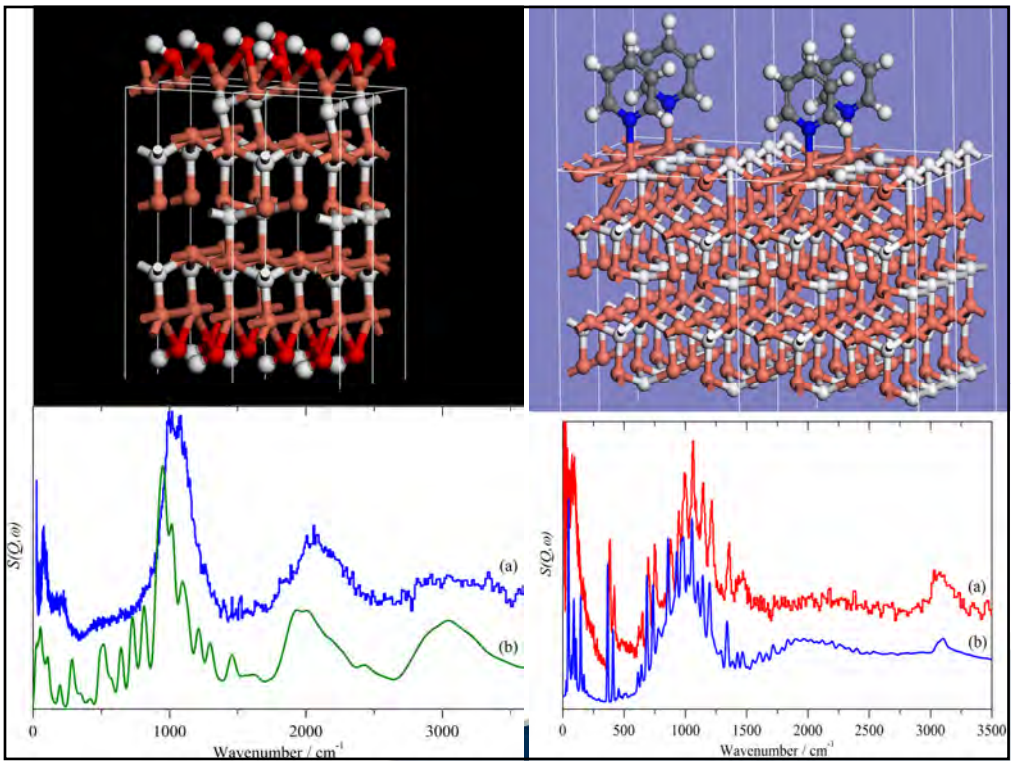
H-C-H bend

0→1	1043
0→2	2054
0→3	2912
O-H	3249

Question 2: How are the products from the aqueous and organic routes different?

CuH/py in the high energy region (MAPS)





CuH summary

For CuH *via* aqueous routes:

A surface layer of hydroxyls is present

For CuH *via* pyridine route:

All pyridine is N coordinated, no evidence for OH.

Difference in the nature of surface species provides a particularly clear example of how an adsorbed layer on a nanoparticle surface determines the properties.

This study was computationally led. *Before* any measurements, the calculations had explained why the IR/Raman data was absent or contradictory and had correctly predicted what the INS spectra should look like.

Modelling explained the differences in behaviour between the aqueous and pyridine routes.

Science & Technology Facilities Council

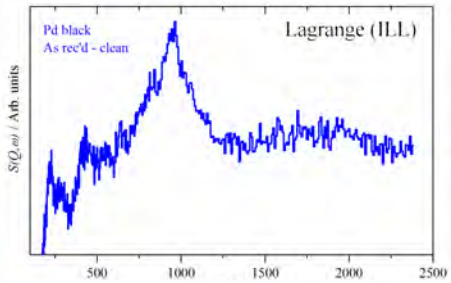
EVONIK INDUSTRIES

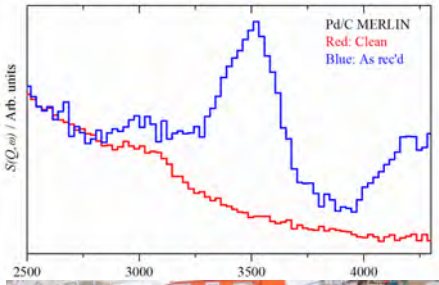
What's on the surface of a precious metal catalyst?

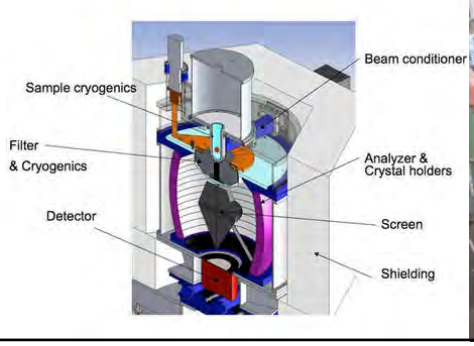
ISIS

Part 1: Pd

Peter Albers (Evonik)



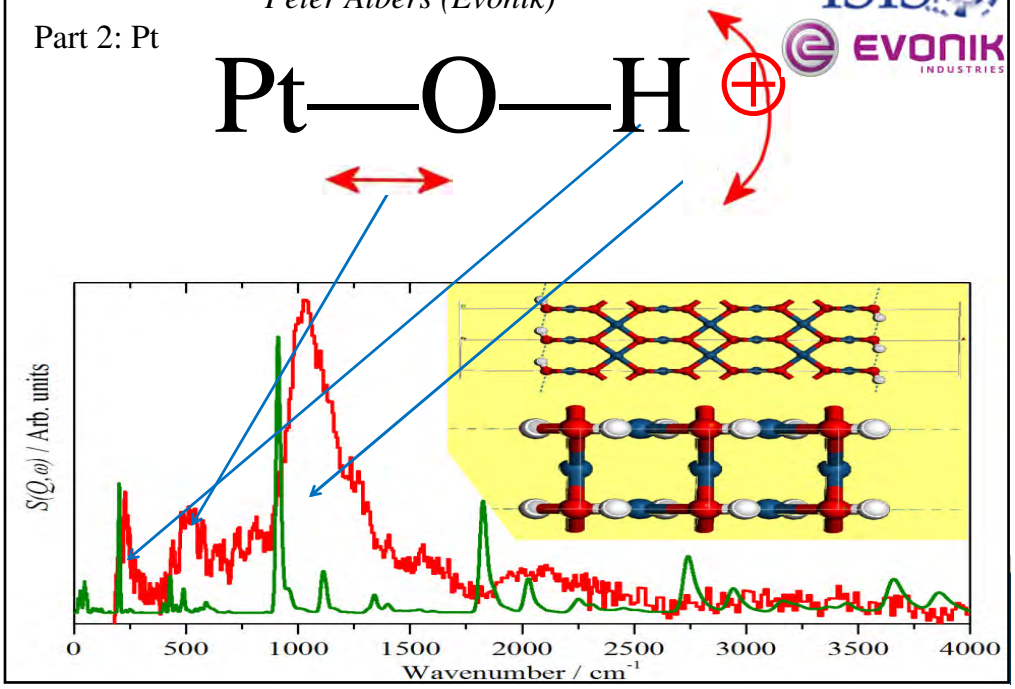


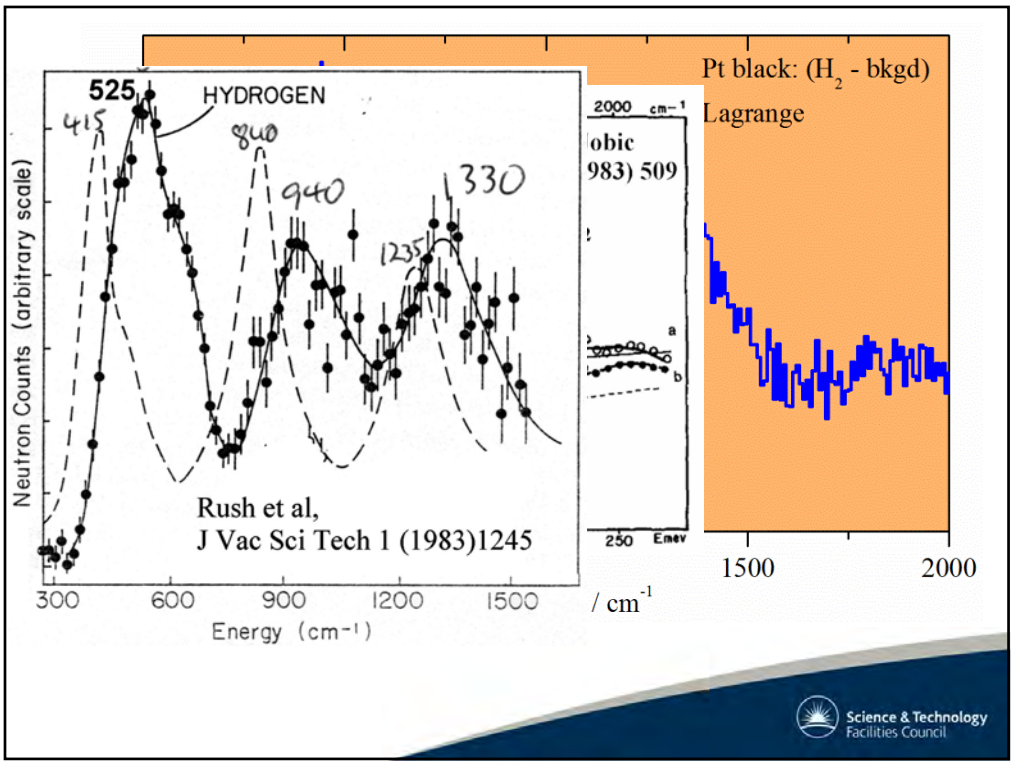


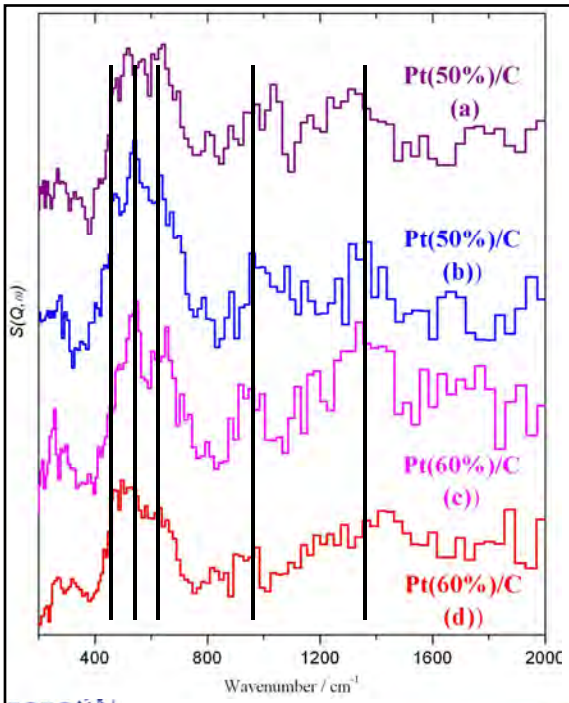
What's on the surface of a precious metal catalyst?

Peter Albers (Evonik)

Part 2: Pt







Adsorbed hydrogen on Pt/C fuel cell catalysts

Different preparations give rise to the same sites, so offers the possibility of tailoring the type of site present

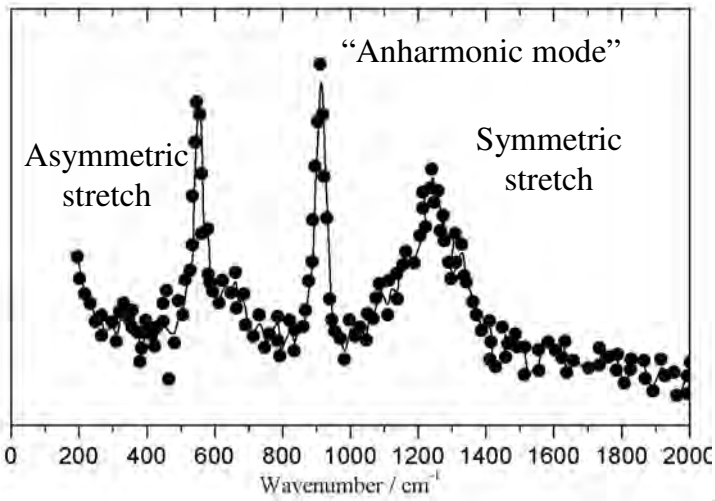
nb: ×10 improvement in S/N by Jan 2017

ISIS

Science & Technology Facilities Council

HREELS of hydrogen on Pt(111)

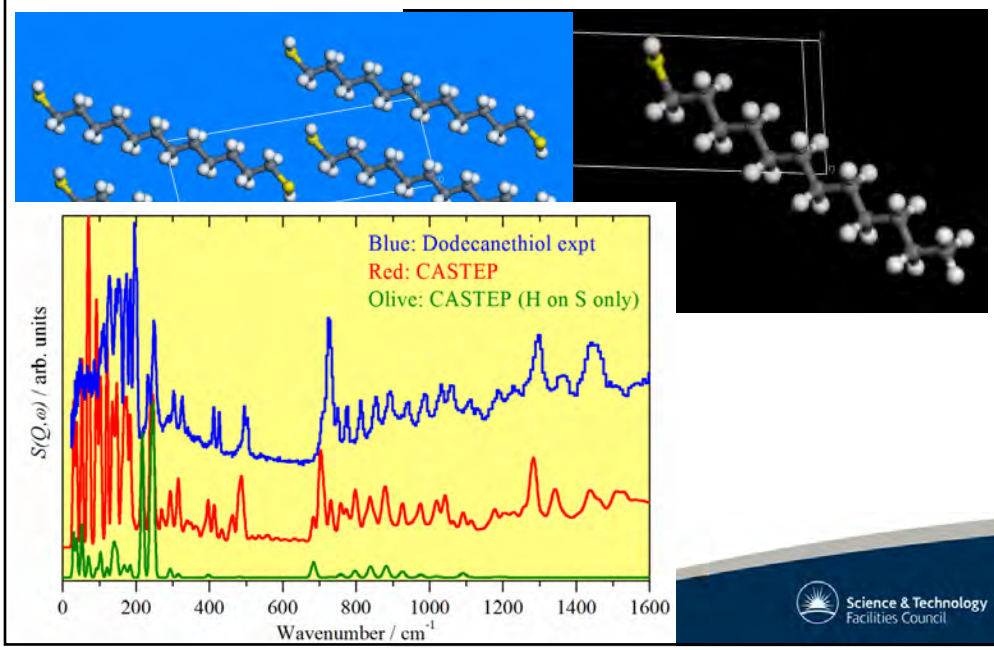
(Bădescu *et al*, Phys. Rev. B 68 (2003) 205401)



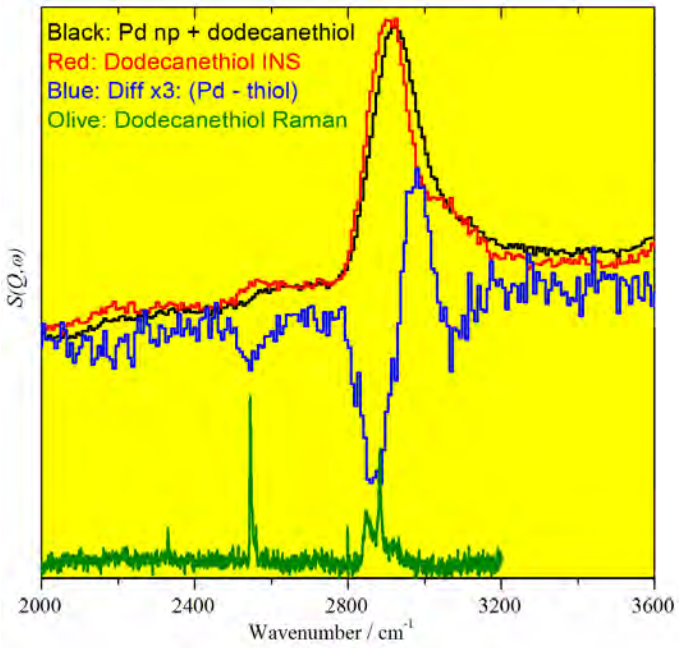
Science & Technology Facilities Council

Dodecanethiol on Pd nanoparticles

Scott Rogers, Nikos Dimitratos, Richard Catlow

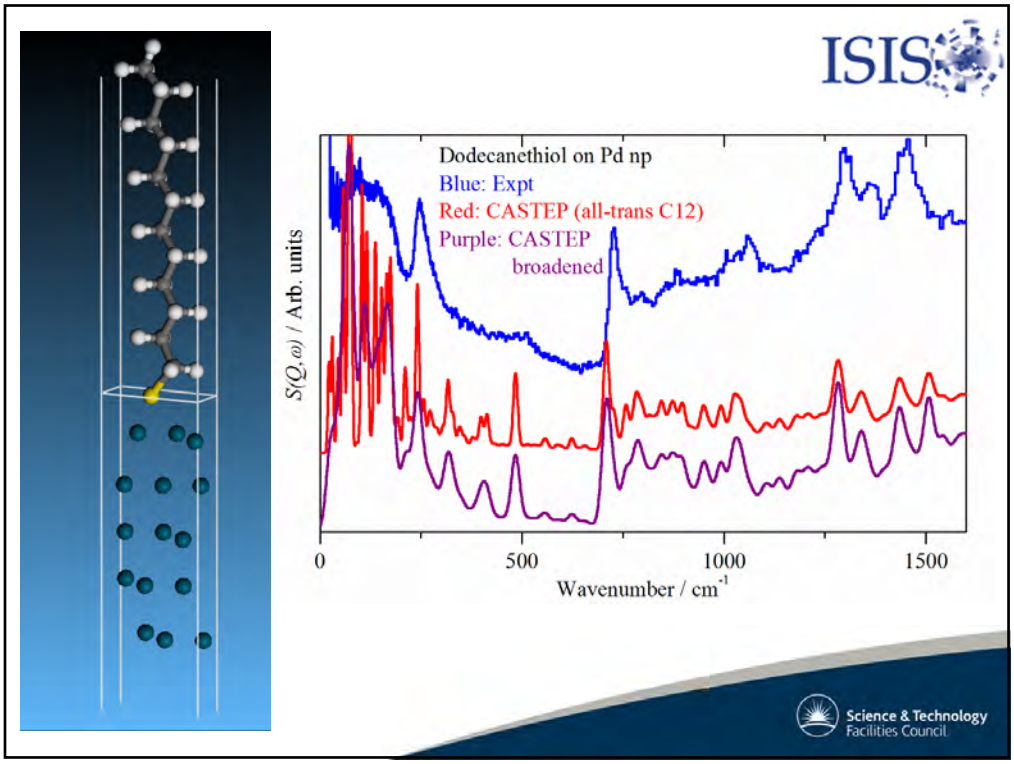


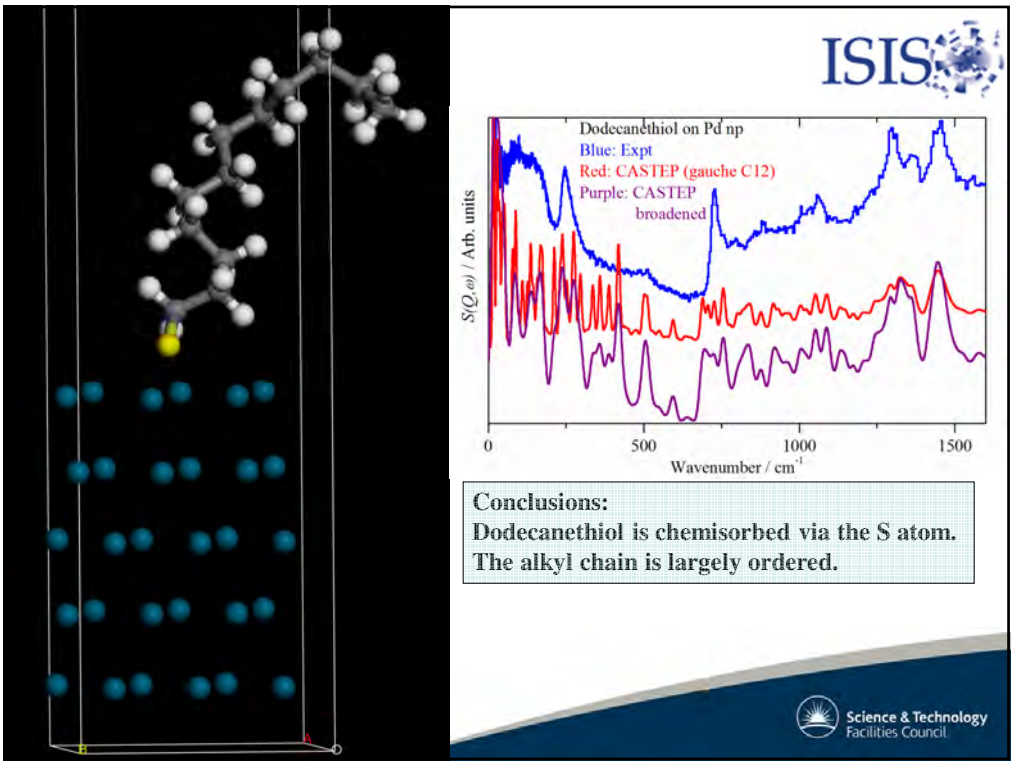
Science & Technology Facilities Council



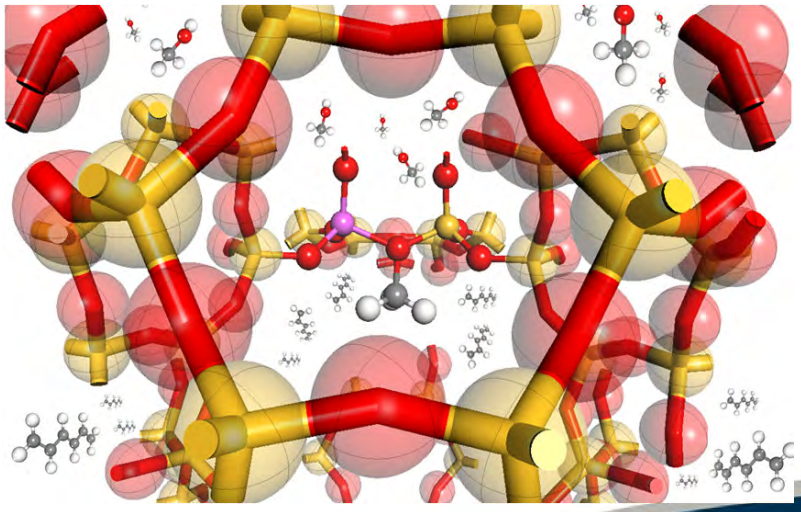
ISIS

Science & Technology Facilities Council


ISIS
Dodecanethiol on Pd np
Blue: Expt
Red: CASTEP (all-trans C12)
Purple: CASTEP
broadened
S(Q,ω) / Arb. units
0
500
1000
1500
Wavenumber / cm-1
Science & Technology
Facilities Council


ISIS
Dodecanethiol on Pd np
Blue: Expt
Red: CASTEP (gauche C12)
Purple: CASTEP
broadened
S(Q,ω) / Arb. units
0
500
1000
1500
Wavenumber / cm-1
Conclusions:
Dodecanethiol is chemisorbed via the S atom.
The alkyl chain is largely ordered.
Science & Technology
Facilities Council

Mechanistic Insight into the Methanol-to-Hydrocarbons Process



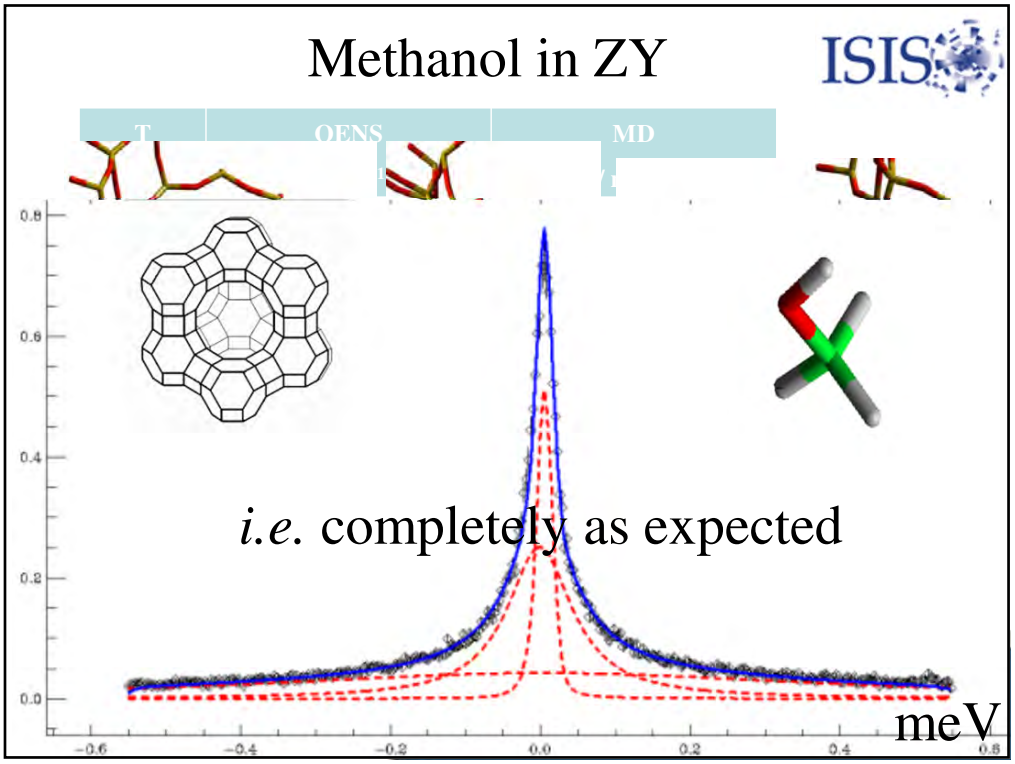
Science & Technology Facilities Council

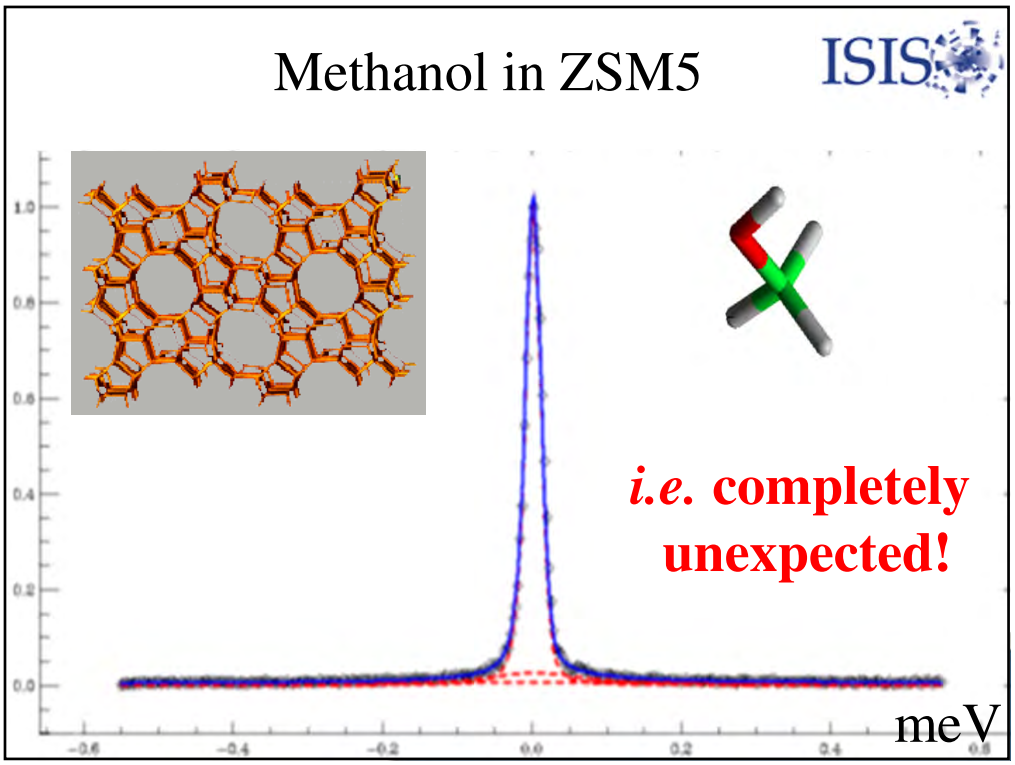
Mechanistic insight into the Methanol-to-Hydrocarbons process

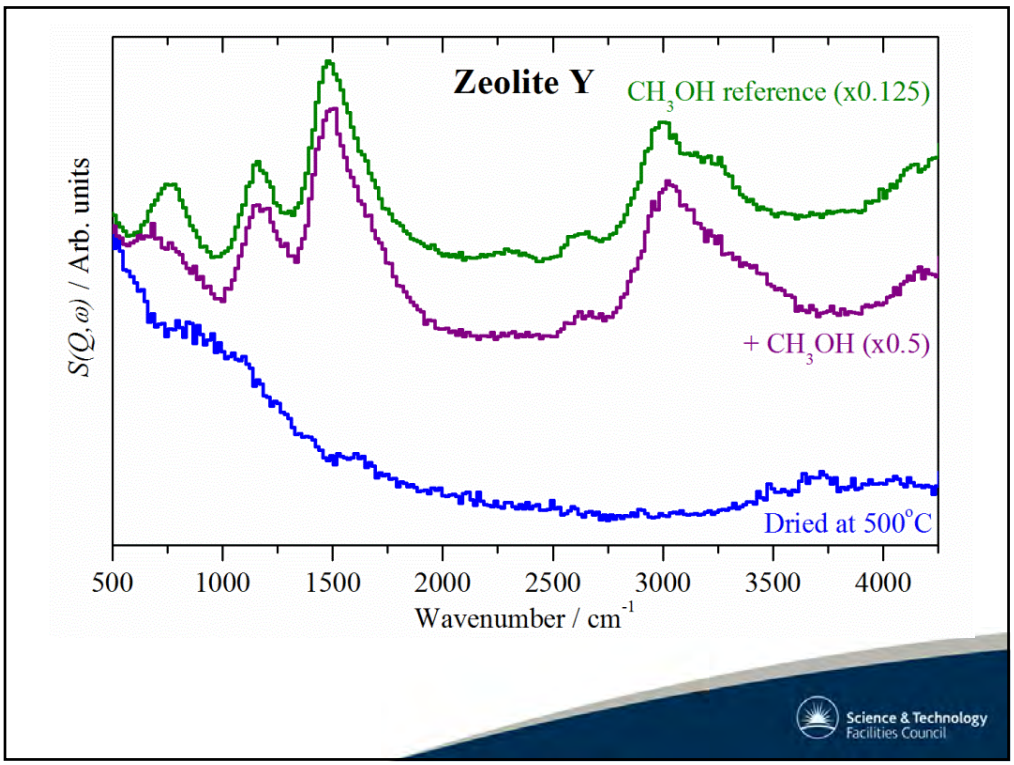
Alex O'Malley, Stewart Parker, Arun Chutia, Ian Silverwood, Vicky Garcia Sakai, Richard Catlow
Chem Comm 52 (2016) 2897-2900.

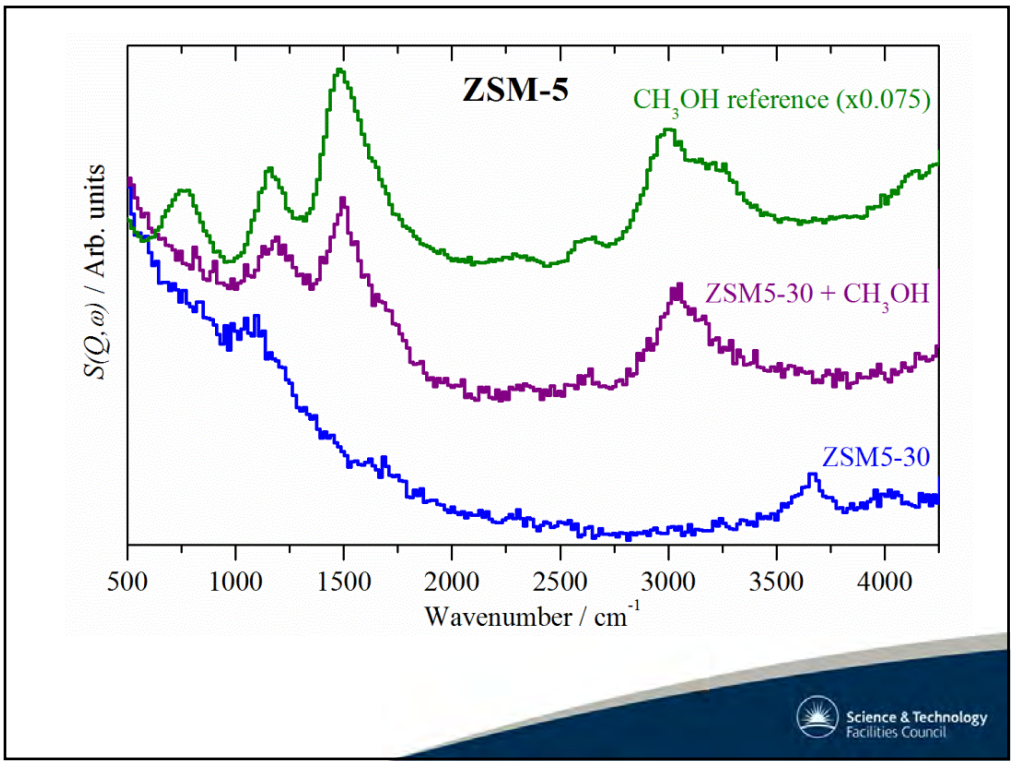
- The diffusion behaviour of methanol in acidic zeolites is of great interest for applications such as the conversion of methanol to gasoline (the MTG process)
- Two acidic catalysts of interest are zeolite-Y (used to convert methanol to dimethylether), and H-ZSM-5.
- The diffusion coefficient of methanol was measured in zeolite HY and H-ZSM-5 using OSIRIS. Measurements were carried out at 300 – 400 K
- The HY QENS data was simulated with molecular dynamics

Science & Technology Facilities Council

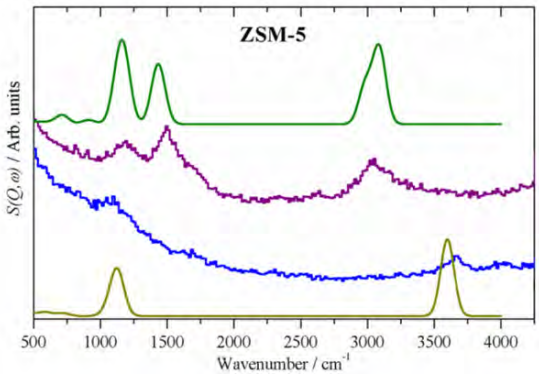

Methanol in ZY
ISIS
T
QENS
MD
0.8
0.6
0.4
0.2
0.0
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
meV
i.e. completely as expected

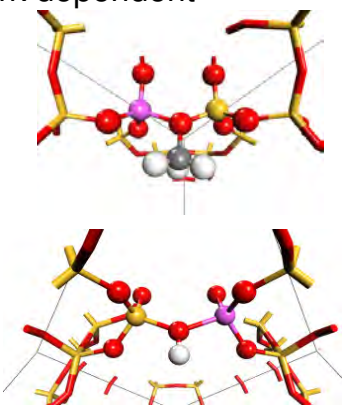

Methanol in ZSM5
ISIS
1.0
0.8
0.6
0.4
0.2
0.0
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
meV
i.e. completely unexpected!


Zeolite Y
CH3OH reference (x0.125)
+ CH3OH (x0.5)
Dried at 500°C
S(Q,ω) / Arb. units
500
1000
1500
2000
2500
3000
3500
4000
Wavenumber / cm-1
Science & Technology Facilities Council


ZSM-5
CH3OH reference (x0.075)
ZSM5-30 + CH3OH
ZSM5-30
S(Q,ω) / Arb. units
500
1000
1500
2000
2500
3000
3500
4000
Wavenumber / cm-1
Science & Technology Facilities Council

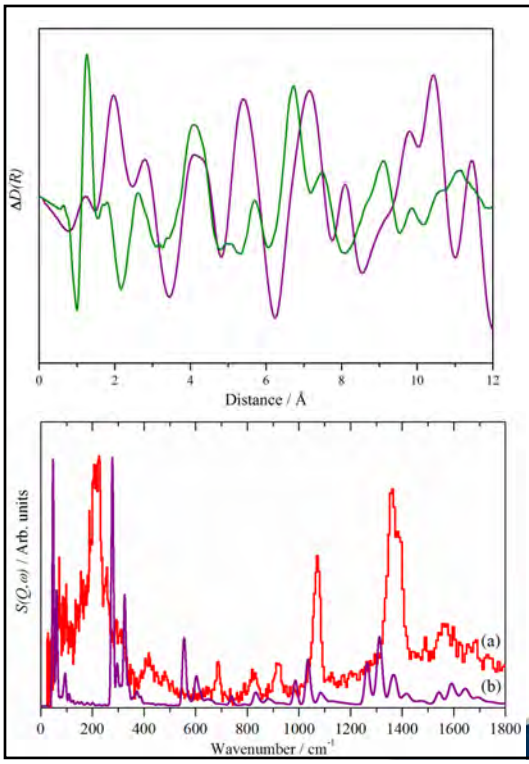
Methoxylation in H-ZSM-5, not HY.
Same Si/Al ratio – Framework dependent





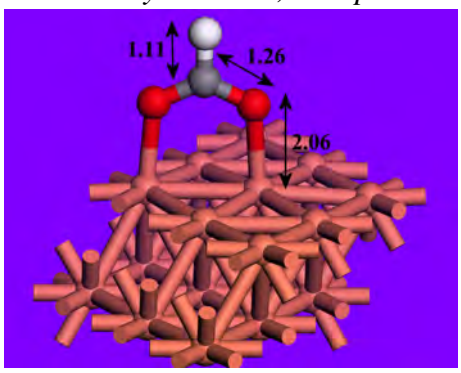
- HY and ZSM-5 have the same Si/Al ratio, but to achieve this HY undergoes steam dealumination which results in weaker Brønsted acid sites
- Excellent agreement achieved for a framework methoxy species in ZSM-5

Science & Technology Facilities Council

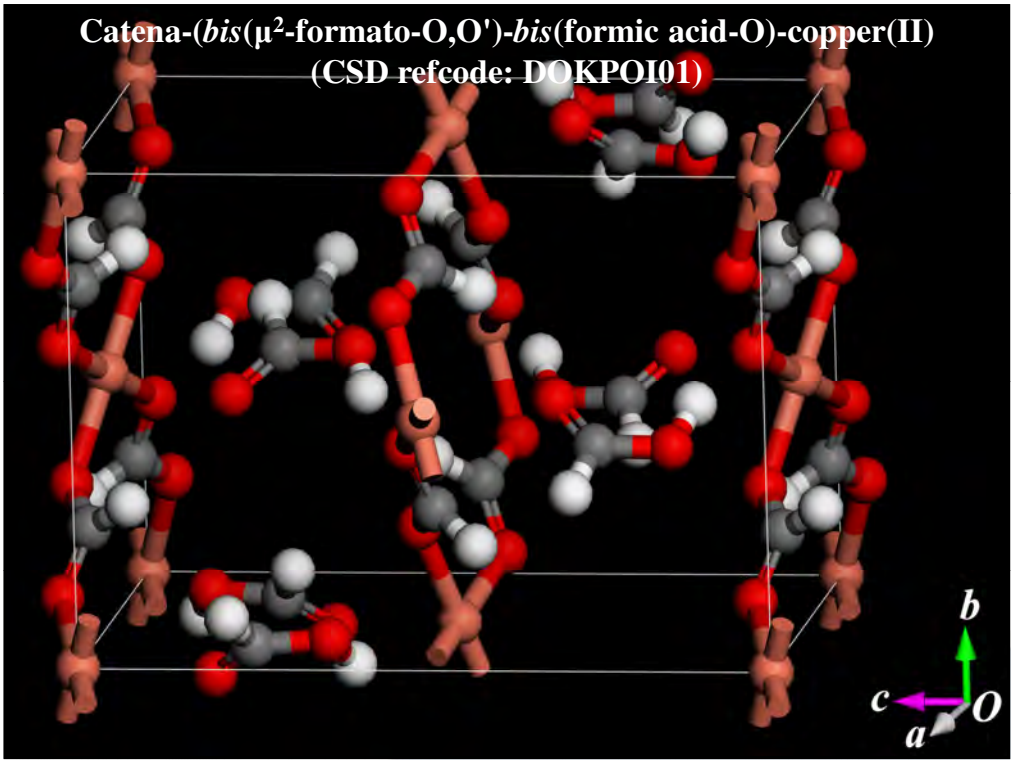


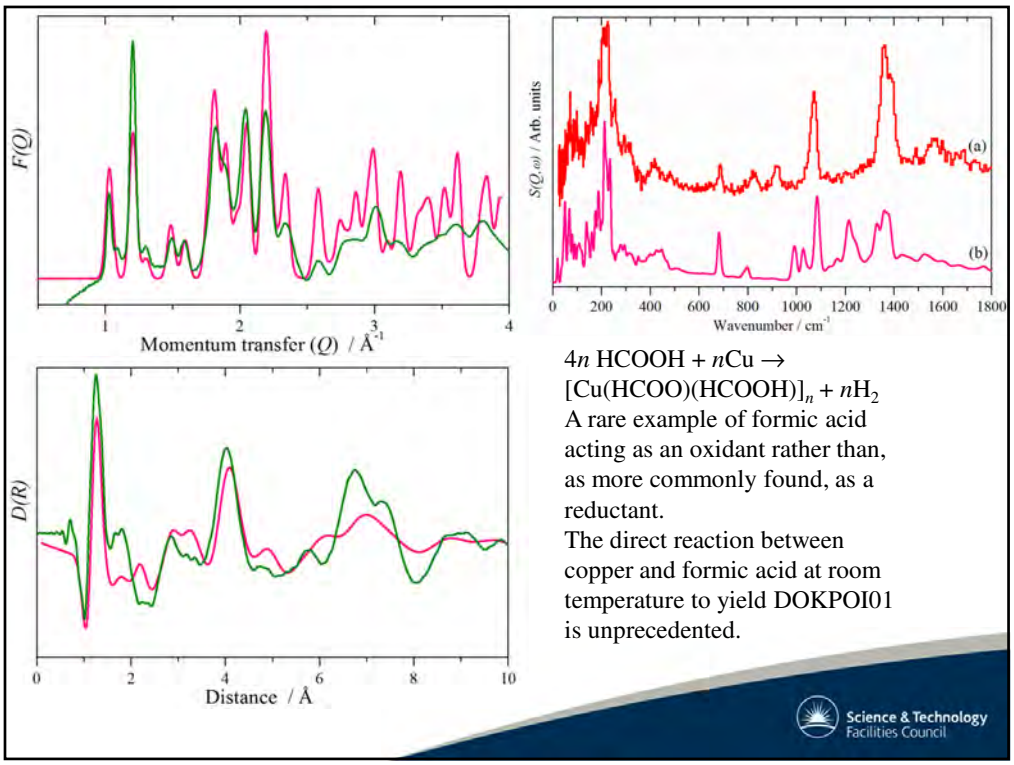
Formate on copper: a cautionary tale!

Callear et al,
Proc. Royal Soc. A, accepted



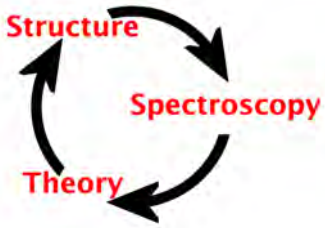
Science & Technology Facilities Council


Catena-(*bis*(μ²-formato-O,O')-*bis*(formic acid-O)-copper(II)
(CSD refcode: DOKPOI01)
b
c
O
a


F(Q)
Momentum transfer (Q) / Å⁻¹
D(R)
Distance / Å
S(Q,ω) / Arb. units
Wavenumber / cm⁻¹
(a)
(b)
4n HCOOH + nCu →
[Cu(HCOO)(HCOOH)]n + nH2
A rare example of formic acid acting as an oxidant rather than, as more commonly found, as a reductant.
The direct reaction between copper and formic acid at room temperature to yield DOKPOI01 is unprecedented.
Science & Technology Facilities Council

Summary

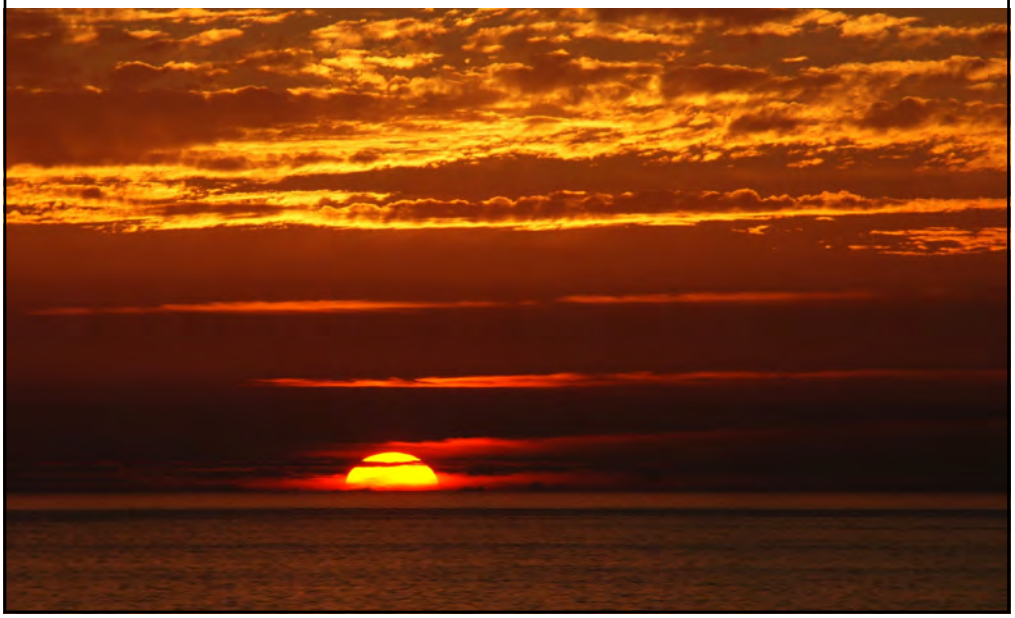
- Neutron scattering in combination with *ab initio* methods enables an in-depth understanding of materials. Systems with long-range order are (usually) tractable, the challenge is to be able to treat disordered and/or nanoparticulate systems with the same rigour.
- Computational methods and neutron scattering are symbiotic: the experimental data provides a stringent test of the calculations, the calculations provide access to properties such as vibrational assignments, electronic structure, diffusion paths that are not readily measurable.
- While MDANSE focusses on analysis of neutron data, the methods are general so can be applied to other types of measurement, *e.g.* structure, infrared, Raman, nmr...
- The range of systems that can be studied is only limited by our imagination!
- Agreement is gratifying, disagreement is more interesting!



Thank you

Science & Technology Facilities Council

Thank you



Force fields and basics of Molecular Dynamics:

An introduction for experimentalists

Miguel A. Gonzalez
Institut Laue-Langevin
Grenoble (France)

Molecular Dynamics

A one slide summary

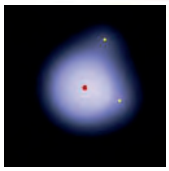
- Start with a set of atomic coordinates:
  - From a crystal structure (database: CIF, PDB, …)
  - From a previous simulation
  - Randomly generated (for a disordered system)
- Give an initial velocity to each atom:
  - From a previous simulation
  - From a random distribution
- Compute the forces (accelerations) acting on each atom

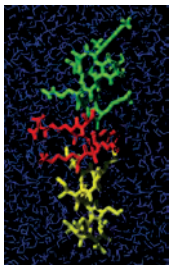
$$m_i \frac{d^2 \vec{r}_i}{dt^2} = -\frac{\partial U(r_1 \ldots r_N)}{\partial \vec{r}_i}$$

- Update the positions

$$\vec{r}_i(t+\Delta t) = \vec{r}_i(t) + \vec{v}_i(t)\cdot\Delta t + \frac{1}{2}\vec{a}_i(t)\cdot\Delta t^2 + \ldots$$

From electrons to forces …





Schrödinger equation

$$\hat{H}\phi(\vec{r},\vec{R}) = E\phi(\vec{r},\vec{R})$$

Born-Oppenheimer approximation

$$\hat{H}_e\phi(\vec{r};\vec{R}) = E_e(\vec{R})\phi_e(\vec{r};\vec{R})$$

Nucleic motion described classically

$$m_i\frac{d^2}{dt^2}\vec{R}_i = -\vec{\nabla}_i E(\vec{R})$$

Need to compromise in accuracy vs efficiency

From Kermode et al., in *Multiscale Simulation Methods in Molecular Sciences*, J. Grotendorst, N. Attig, S. Blügel, D. Marx (Eds.), NIC Series, Vol. 42, pp. 215-228 (2009).

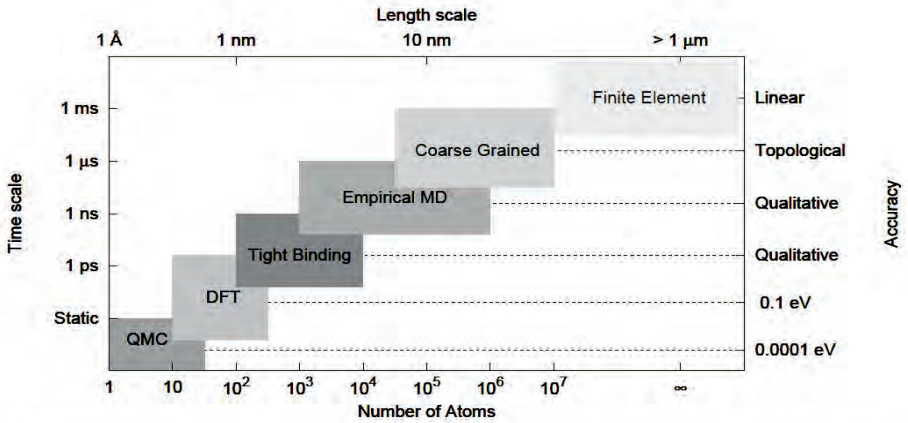
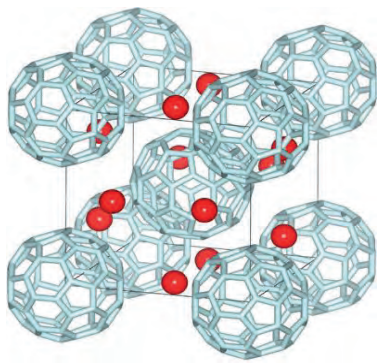
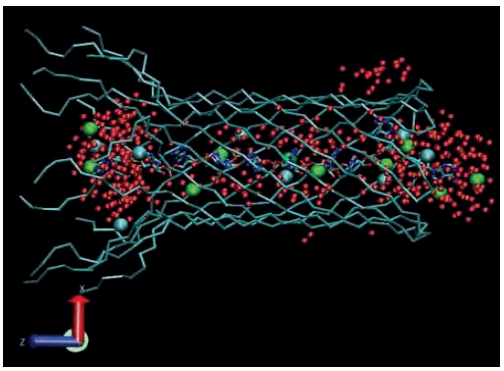


Figure 1. Schematic representation of the range of length- and time-scales accessible to a variety of modelling methods, from quantum Monte Carlo (QMC) for very accurate, very expensive static calculations through to approximate methods such as finite-element modelling.

Molecular Mechanics (MM)

- No electrons
- Molecules represented as 'balls' + 'springs'
- Empirical potential to account for intra- and intermolecular interactions
  →**Force Field (FF)**
- ~5-6 orders of magnitude faster than DFT → ~10^4-10^5 atoms, ~1-100 ns

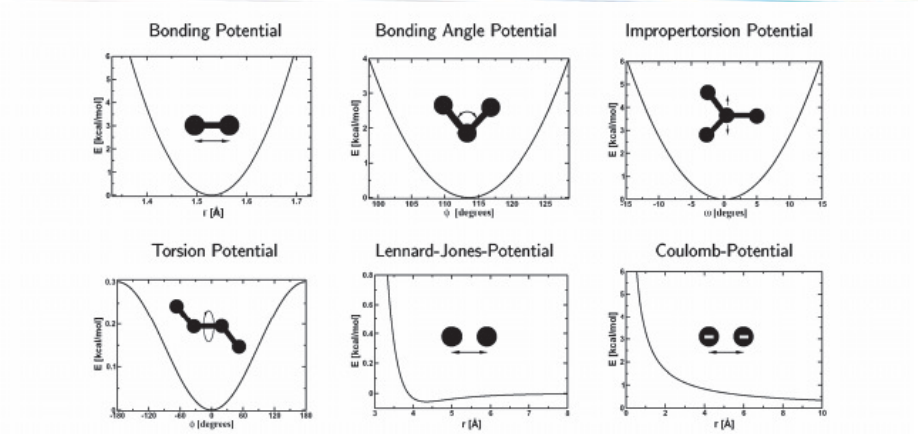




What is a Force Field?

- A mathematical expression describing the dependence of the energy of a system on the coordinates of the atoms.
- Consists of:
  1. An analytical form of the interatomic potential energy, $U(r_1 \ldots r_N)$
  2. A set of parameters entering $U(r_1 \ldots r_N)$
- The parameters are obtained from *ab initio* calculations and/or fitting to experimental data.
- Molecules are defined as collection of atoms held together by simple elastic (harmonic) forces.

A standard FF

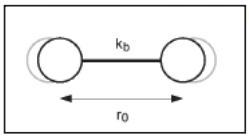


$$E = \sum_b k_b (r - r_b)^2 + \sum_\theta k_\theta (\theta - \theta_0)^2 + \sum_\omega k_\omega (\omega - \omega_0)^2$$
$$+ \sum_\phi k_\phi (1 - \cos(n\phi - \delta)) + \sum_{i<j} 4\,\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6} \right] + \sum_{i<j} \frac{q_i q_j}{r}$$

Intramolecular terms: Bonding

- Very often, represented by a simple harmonic function:

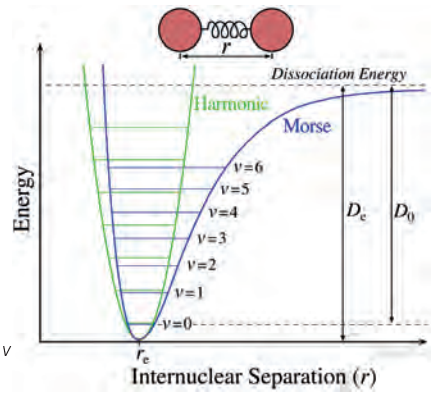
$$V_b(r_{ij}) = \frac{k_b}{2}(r_{ij} - r_0)^2$$



- This is a poor approximation at large values of r_{ij}, but at normal temperatures r_{ij} is small (stiff bonds).

- Other functional forms used occasionally, in particular we can employ the Morse potential to improve the accuracy (at the cost of slowing the calculation).

$$V(r) = D_e[1 - e^{-a(r - r_0)}]^2$$

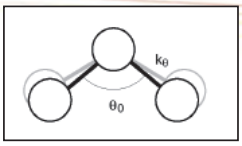


Two new parameters: D_e: dissociation energy
a: width of the potential well

Intramolecular terms: Bending

• Again represented usually by a harmonic function:

$$V_\theta(\theta_{ijk}) = \frac{k_\theta}{2}(\theta_{ijk} - \theta_0)^2$$



• Bending constants, k_θ , are normally smaller than stretching constants k_b by a factor 5 to 10.

• In CHARMM FF, there is an additional term, called Urey-Bradley, introduced to optimize the fit to vibrational spectra:

$$V_{UB} = \sum_{UB} \frac{1}{2} k_{UB}(s - s_0)^2$$

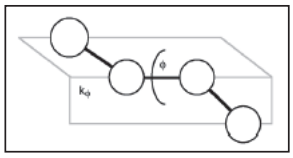
s: distance between 1,3 atoms (separated by 2 bonds)

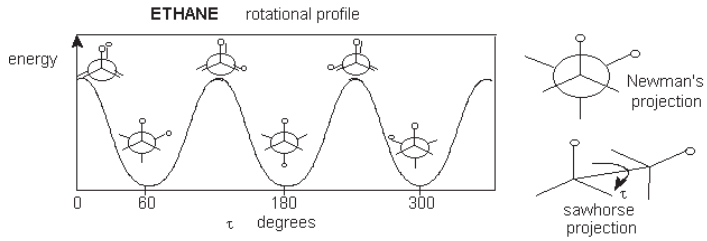
Intramolecular terms: Torsion

• Typically represented by cosine functions:

$$V_\varphi(\varphi) = \frac{k_1}{2}(1+\cos\varphi) + \frac{k_2}{2}(1+\cos 2\varphi) + \frac{k_3}{2}(1+\cos 3\varphi)$$

$$U_{torsion} = \sum_{torsions} \frac{V_n}{2}[1+\cos(n\phi - \delta)]$$



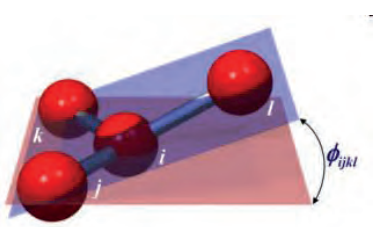


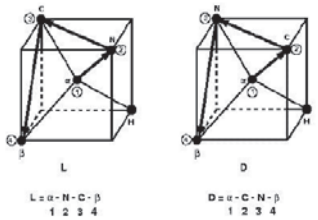
• Torsional motions are typically hundreds of times less stiff than bond stretch.
• They ensure correct degree of chain rigidity.
• Necessary to reproduce major conformational changes.

Intramolecular terms: Improper torsion

- Used to describe the energy of out-of-plane motions.
- It is often necessary for planar groups, such as sp^2 hybridized carbons in carbonyl groups and in aromatic rings, because the normal torsion terms described above is not sufficient to maintain the planarity.
- Also needed sometimes to define chiral molecules.

$$U_{improper} = \sum_{improper} \frac{V_2}{2}[1+\cos(2\omega - 180^\circ)] \qquad U_{improper} = \sum_{improper} \frac{k_w}{2}(\omega - \omega_0)^2$$





Intermolecular terms: Van der Waals

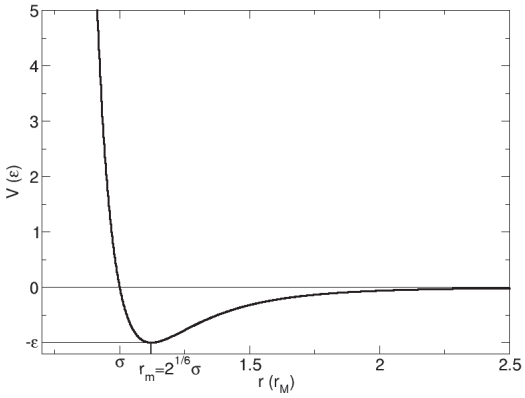
- Van der Waals interactions between two atoms arise from the balance between repulsive and attractive forces:
  - Repulsion is due to overlap of electron clouds of two atoms.
  - Attraction is due to the dispersion or induced multipole interactions (dominated by induced dipole-dipole interactions).
- Very often they are represented by the Lennard-Jones potential.
- We will have VdW interactions between all atoms that are not directly bonded: atoms belonging to different molecules but also between atoms of same molecule but separated (often, but not always) by more than 3 bonds (> 1-4 interactions).

VDW: Lennard-Jones

$$V_{LJ} = \sum_{i<j} 4\varepsilon_{ij}\left[\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right]$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \qquad \sigma_{ij} = \frac{1}{2}(\sigma_i + \sigma_j)$$

$$\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$$



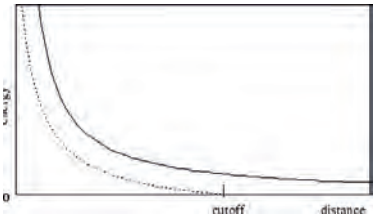
- Calculated pairwise but parameterized against experimental data → includes implicitly many-body effects.
- Use of r^{-12} for the repulsive term due to computational convenience.
- Other functions in use, but less popular because of higher CPU cost:

$$V(r) = ae^{-br} - \frac{c}{r^6}$$ (less repulsive, some inorganic systems)

Intermolecular terms: Electrostatics

- Electrostatic interactions between charged atoms and dipolar groups are normally treated by assigning a point charge (**partial charge**) to each atom and then using Coulomb's law:

$$V_{\text{elec}}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$



- In standard models, the charges assigned are **effective charges**, i.e. they include polarization effects appearing in condensed phases (e.g. for water, μ_{models} = 2.3-2.7 D while $\mu_{H2O,\ gas}$ = 1.8 D).

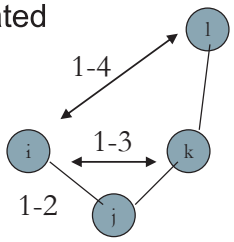
- **Long-ranged**: decay as 1/r (charge-charge) or $1/r^3$ (dipole-dipole) → ***Care with truncation!***

Non-bonding exclusions

Partition of potential energy in different terms is arbitrary!

Which intramolecular pairs should we include in the sum of VdW and electrostatics energies?

- Possible problems:
  - Small r_{ij} → large non-bonding interactions
  - Already bonding terms (bonds, angles, dihedrals) between them
- **Non-bonded exclusions:** VdW and Coulomb interactions calculated only for pairs not involved in direct bonding interactions.
- Number and type depend on the FF. Typically:
  - 1-2 and 1-3 interactions excluded
  - 1-4 interactions partially excluded
- Combination of dihedral + 1-4 VdW & electrostatics will determine rotational barriers → **care if mixing parameters from different FFs!**



Special terms

Not commonly used and not available in all MD software

$$U_{total} = U_{intramol} + U_{cross\text{-}term} + U_{non\text{-}bond}$$

$$U_{cross\text{-}term} = U_{bond\text{-}bond} + U_{bond\text{-}bend\text{-}bond} + U_{bend\text{-}bend} + U_{dihedral\text{-}bond} + \ldots$$

$$U_{non\text{-}bond} = U_{vdW} + U_{qq} + U_{pol} + U_{HB}$$

U_{cross} brings corrections to energy due to couplings between stretches, bends, and torsions, e.g:

$$U_{bond-bond} = \frac{1}{2} k_{bb}(r - r_0)(r' - r_0') \qquad U_{bond-bend} = \frac{1}{2} k_{ba}[(r - r_0) + (r' - r'_0)](\theta - \theta_0)$$

U_{HB} may improve the accuracy of H-bonding energy. $U_{HB}(r) = \frac{A}{r^{12}} - \frac{C}{r^{10}}$

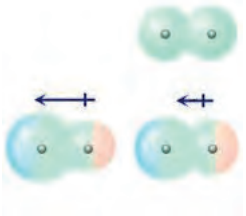
Adding the effect of polarizability

• Surrounding molecules will induce a charge redistribution.

• This can be modelled using fluctuating charges, shell models (Drude particle) or induced point dipoles.

$$U_{pol} = -\frac{1}{2}\sum_i \mu_i \boldsymbol{E}_i^0$$

$$\mu_i = \alpha_i \boldsymbol{E}_i = \alpha_i \left[\boldsymbol{E}_i^0 + \sum_{j \neq i}^{N} \boldsymbol{T}_{ij} \mu_j \right]$$



• Iterative calculation:
- Evaluate induced dipoles due to permanent charges.
- Adjust charges or reevaluate induced dipoles due to $q+\mu^{ind}$.
- Repeat until convergence.
- Reiterate every time one molecule is moved.

Polarizable FFs

• Very important to increase transferability.

• But FF is a global entity → adding polarizability needs re-optimizing VDW and intramolecular parameters.

• As a consequence they are not so matured as non-polarizable and in some cases (e.g. water) results are not better.

• But it is agreed that they will be able to provide better agreement with experiment than non-polarizable FF (Cieplak et al., *J. Phys.: Condens. Matter* **21**, 333102 (2009)).

• Polarizable FFs: PIPF, DRF90, AMOEBA, XPOL + some of the classical FFs (CHARMM, AMBER, OPLS, GROMOS).

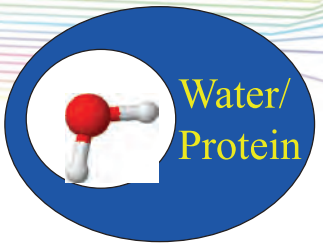
FF parameterization

An elaborate job

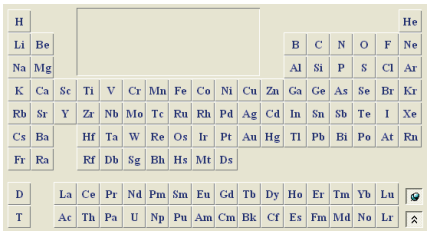
- Goal: Describe in classical terms the quantum mechanical facts by partitioning the total electronic energy into well separated atom-atom contributions.
- As it is impossible to fully separate the intricate electronic effects, this means applying significant approximations → empirical potentials.
- We will obtain different FFs depending on the set of data (experimental or QM calculations) and the procedure employed to optimize the parameters.
- "**FF development is still as much a matter of art as of science**" (T. Halgren, *Curr. Opin. Struct. Biol.*, 1995)

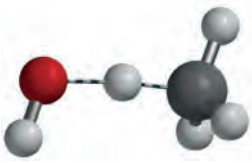
FF development

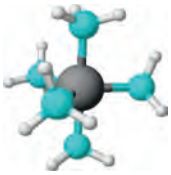
Current topics



- Explicit polarizable force fields
- Force fields for inorganic compounds (especially metals)
- Universal force fields
- Reactive force fields
- Force matching







FF parameterization

Components

- Functional forms of the components of the energy expression
- A list of atom types
- A list of atomic charges
- Rules for atom types
- A set of parameters for the function terms
- (Rules to generate parameters that have not been defined explicitly)

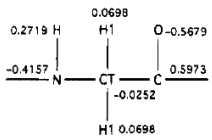
Generally developed by specialized groups and available in the literature as data tables

FF example

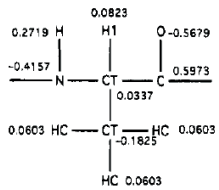
Amber

Table 1. List of Atom Types[a]

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 at 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^2 nitrogen in amides
	NA	sp^2 nitrogen in aromatic rings with hydrogen attached (e.g. protonated His, Gua, Trp)
	NB	sp^2 nitrogen in 5-membered ring with lone pair (e.g. N7 in purines)
	NC	sp^2 nitrogen in 6-membered ring with lone pair (e.g. N3 in purines)
	N*	sp^2 nitrogen in 5-membered ring with carbon substituent (in purine nucleosides)
	N2	sp^2 nitrogen of aromatic amines and guanidinium ions
	N3	sp^3 nitrogen



GLY



ALA

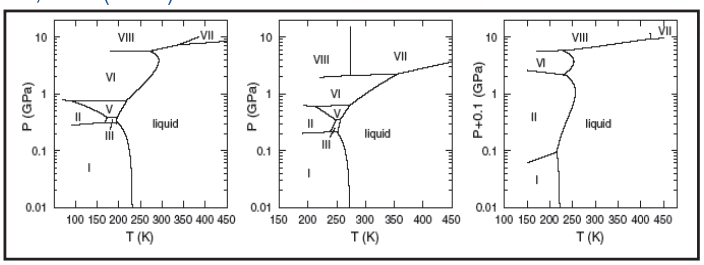
Van der Waals Parameters

atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k
C^l	1.9080	0.0860	H2	1.2870	0.0157	HS	0.6000	0.0157	O2	1.6612	0.2100
CA	1.9080	0.0860	H3	1.1870	0.0157	HW	0.0000	0.0000	OH	1.7210	0.2104
CM	1.9080	0.0860	H4	1.4090	0.0150	IP	1.8680	0.00277	OS	1.6837	0.1700
Cs	3.3950	0.0000806	H5	1.3590	0.0150	K	2.6580	0.000328	OW	1.7683	0.1520
CT	1.9080	0.1094	HA	1.4590	0.0150	Li	1.1370	0.0183	P	2.1000	0.2000
F	1.75	0.061	HC	1.4870	0.0157	N^m	1.8240	0.1700	Rb	2.9560	0.00017
H	0.6000	0.0157	HO	0.0000	0.0000	$N3^n$	1.875	0.1700	S	2.0000	0.2500
H1	1.3870	0.0157	HP	1.1000	0.0157	O	1.6612	0.2100	SH	2.0000	0.2500

Choosing a FF

Highly subjective, e.g. water case

Sanz, PRL (2004)



TIP4P **Exp** **SPC/E**

But ...

$D_{SPC/E}$ = 2.49, D_{TIP4P} = 3.29, D_{exp} = 2.4 (10^{-5} cm^2/s)

$\varepsilon_{SPC/E}$ = 71, ε_{TIP4P} = 53, ε_{exp} = 78

Table. Main water models of the literature developed since 1933. In the following R is for rigid, F for flexible, D for dissociable and P for polarizable. Values of the dipole moment in the gaseous phase and in the liquid phase are indicated by μ_g and μ_l, respectively.

Acronym (date)	Reference	Status	Type	Sites	$\mu_g(D)$	$\mu_l(D)$
BF (1933)	7	empirical	R	4	2.0	2.0
R (1951)	21	empirical	R	5	1.84	1.84
BNS (1971)	22	empirical	R	5	2.17	2.17
ST2 (1973,1993)	23	empirical	R	5	2.35	2.35
CF (1975,1978,1995)	24	empirical	F	3	1.86	1.98
MCY (1976)	25	ab initio	R	4	2.19	2.19
DCF (1978,1980,1993)	26	empirical	F,D,P	3	1.855	-
PE (1979)	27	empirical	P	1	1.855	2.50
SPC (1981)	14	empirical	R	3	2.27	2.27
TIP3P (1981,1983)	15	empirical	R	3	2.35	2.35
RWK (1982)	28	empirical	F	4	1.85	1.89
TIP4P (1983)	15	empirical	R	4	2.18	2.18
BJH (1983)	29	empirical	F	3	1.87	1.99
SPC/F (1985)	30	empirical	F	3	2.27	2.42
MCYL (1986)	31	ab initio	F	4	2.19	2.26
SPC/E (1987)	32	empirical	R	3	2.35	2.35
WK (1989)	33	empirical	R	4	2.60	2.60
SPCP (1989)	34	empirical	P	3	1.85	2.90
CKL (1990)	35	empirical	F,P	4	1.88	2.20
MCHO (1990)	36	ab initio	P	6	2.12	≈3.0
NCC (1990)	37	ab initio	P	6	1.85	2.80
NEMO (1990,1995)	38	ab initio	P	5	2.04	2.89
PTIP4P (1991)	39	empirical	P	4	1.85	2.80
SPC/FP (1991)	40	empirical	F,P	3	1.85	2.44
PSRWK (1991)	41	empirical	P	4	1.88	2.63
KJ (1992)	42	empirical	P	4	1.85	-
NCCvib (1992)	37b	ab initio	F,P	6	1.85	3.11
ASP-W (1992,1998)	43	ab initio	P	3	1.85	2.90
RPOL (1992)	44	empirical	P	3	2.02	2.62
CPMD (1993,1999)	45	DFT+CP	F,D,P	nucl.+el.	1.87	2.95
PPC (1994)	46	ab initio	P	4	2.14	2.51
SPC/FQ (1994)	47	empirical	P	3	1.85	2.83
TIP4P/FQ (1994)	47	empirical	P	4	1.85	2.62
KKY (1994)	48	empirical	F,D	3	2.38	2.21
SQPM (1995)	49	valence bond	P	4	1.85	2.62
SCPDP (1996)	50	empirical	P	4	1.85	2.87
TAB/10D (1998)	51	SCF+MD	P	5	1.85	2.65
NSPCE (1998)	52	empirical	R	3	2.18	2.18
NCF (1998)	53	empirical	F	3	1.85	1.90
MCDHO (2000)	54	ab initio	F, P	4	1.85	3.01
TIP5P (2000)	55	empirical	R	5	2.29	2.29
SPC/HW (2001)	56	empirical	R	3	2.41	2.41
DEC (2001)	57	empirical	R	3	1.85	1.85
SWFLEX (2001)	58	empirical	P	4	1.85	2.59
POLARFLEX (2001)	59	valence bond	F,P	3	1.85	2.55
POL5 (2001)	60	ab initio	P	5	1.85	2.71

Guillot, J. Mol. Liq. (2002)

Choosing a FF

Highly subjective, e.g. water case

- Compute 17 different properties and note each of them.
- TIP4P/2005 is probably close to the best that can be achieved with a rigid non-polarizable model.

Vega & Abascal, *Phys. Chem. Chem. Phys.* **13**, 19663-19688(2011)

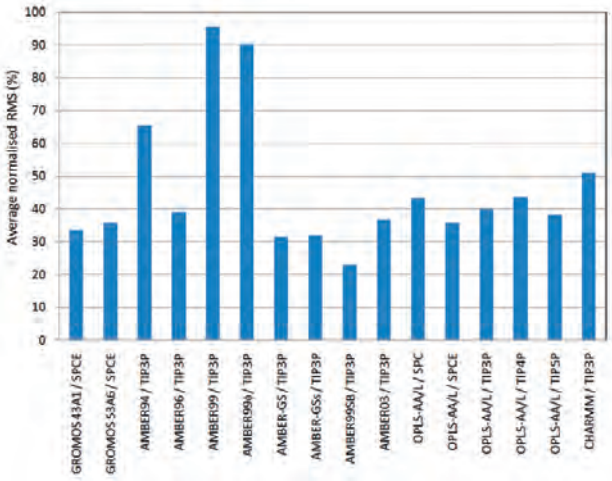
Table 4 Scoring summary

Property	TIP3P	SPC/E	TIP4P	TIP4P/2005	TIP5P
Enthalpy of phase change	4.0	2.5	7.5	5.0	8.0
Critical point properties	3.7	5.3	6.3	7.3	3.3
Surface tension	0.0	4.5	1.5	9.0	0.0
Melting properties	2.0	5.0	6.3	8.8	4.5
Orthobaric densities and TMD	1.8	5.5	4.0	8.5	4.0
Isothermal compressibility	2.5	7.5	2.5	9.0	4.0
Gas properties	2.7	0.7	1.3	0.0	1.0
Heat capacity at constant pressure	4.5	3.5	4.0	3.5	0.0
Static dielectric constant	2.0	2.3	2.3	2.7	2.3
T_m–TMD–T_c ratios	3.7	6.7	8.7	8.3	6.7
Densities of ice polymorphs	3.5	5.0	6.0	8.8	2.3
EOS high pressure	7.5	8.0	7.5	10	5.5
Self-diffusion coefficient	0.3	8.0	4.3	8.0	4.5
Shear viscosity	1.0	7.5	2.5	9.5	4.0
Orientational relaxation time	0.0	6.0	2.0	9.0	4.0
Structure	4.0	6.0	6.0	7.5	7.5
Phase diagram	2.0	2.0	8.0	8.0	2.0
Final score	2.7	5.1	4.7	7.2	3.7

Choosing a FF

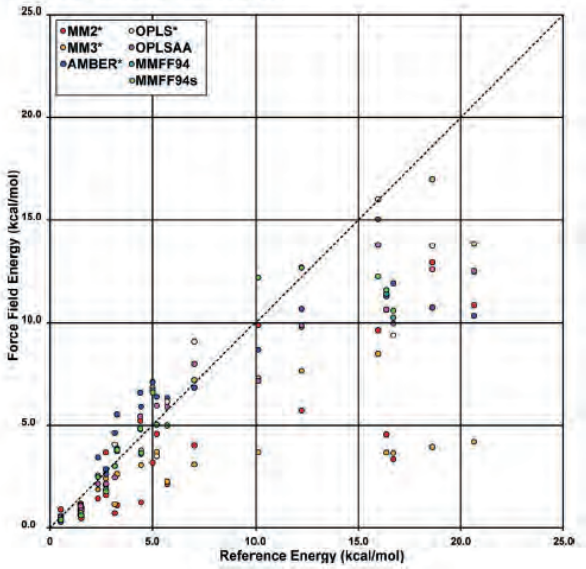
Highly subjective

Paton & Goodman, *J. Chem. Inf. Model* **49**, 944-955 (2009)



Aliev & Courtier-Murias, *JPCB* **114**, 12358-12375 (2010)

Chart 1. Comparison of Force Field Interaction Energy with Ab Initio Values for the S22 Data Set[a]



Choosing a FF

Highly subjective

"All leading force fields now provide reasonable results for a wide range of properties of isolated molecules, pure liquids, and aqueous solutions", Jorgensen & Tirado-Rives, *PNAS* **102**, 6665 (2005).

But the best force field does not exist!

Choice will depend on the type of molecular system and the type of properties we are interested in!

Force fields

Summary

- The FF must be **simple** enough to be evaluated quickly, but sufficiently detailed to **reproduce the main features** of the system.
- They are constructed by parameterising a potential function using either **experimental** data (n, X-R and e^- diffraction, NMR, IR, Raman and neutron spectroscopy) **or ab-initio and semi-empirical** quantum mechanical calculations.
- The goal is to replace the true potential with a simplified model **valid in the region being simulated**.
- A force-field is a single entity, where the parameters are optimized in a global way. Therefore they are not strictly independent and some **caution is needed before modifying one parameter or mixing parameters** from different force-fields.

FF vs DFT

Disadvantages

- **Intrinsic limitations:**
  - No electronic structure information is available.
  - Unable to handle reactions (bond-breaking/forming, charge transfer, ...)
- **Limited prediction power:**
  - The accuracy depends on the parameterization.
  - It can only be used for systems having the functional groups that were included in parameterization.
  - Limited transferability of the FF (care needed when applying a FF under conditions that are very different from the conditions used in its parameterization, e.g. P, T)

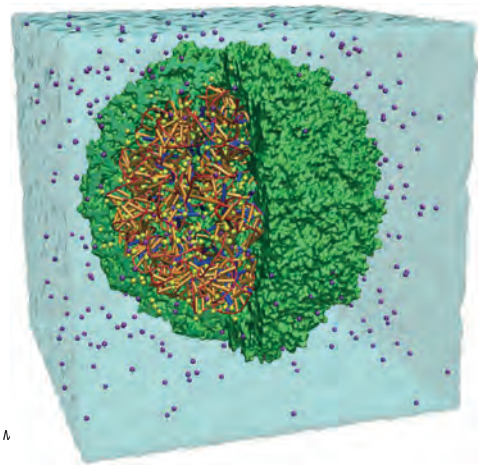
Needs experimental validation!

FF vs DFT

Advantages

Allow to handle large systems and simulate relatively long times (several orders of magnitude faster):

- Small cluster (16 processors) ~ 10^4 atoms @ 1 ns/day
- BlueGene/L (131072 processors) ~ 320 billion atoms (a cubic piece of metal of side ~1 μm) @ 10 ps/day (Kadau, *Int. J. Modern Physics C* (2006))
- Anton (spec design for MD) ~ 10^4 atoms @ 10 μs/day (Klepeis, *Curr. Opin. Str. Biol.* (2009))
- NCS (256 nodes) ~ 10^6 atoms (all-atom satellite tobacco mosaic virus, NAMD) @ 1ns/day (Freddolino, *Structure* (2006))
- Perspectives: multimillion-atom (~100 nm scale) @ 30 ns/day on a Cray XT5 (10^5 cores) (Schulz, *J. Chem. Theory Comput.* (2009))



FF vs DFT

Advantages

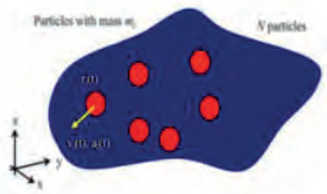
- **Analysis of energy contributions can be done at the level of individual interactions or classes of interactions.**

- **It is possible to modify the energy expression to bias the calculation.**

- **But note that the energy partition in different terms (intramolecular, VDW, electrostatic, …) does not have a precise physical meaning!**

Molecular Dynamics

Generate a dynamical trajectory by integrating Newton's equations of motion, with suitable initial and boundary conditions. We need a good way to determine the forces acting on each atom (e.g. FF or DFT) and a accurate numerical method to integrate the equations.



F = ma

$$m\frac{d^2 r_j}{dt^2} = -\nabla_{r_j} U(r_j) \qquad j = 1..N$$

Total energy of system

$$E = K + U$$

$$K = \frac{1}{2} m \sum_{j=1}^{N} v_j^2$$

$$U = U(r_j)$$

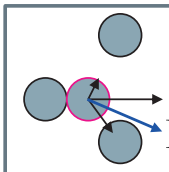
Coupled system N-body problem, no exact solution for N>2

Molecular Dynamics

Integrating the equations of motion

$$\frac{d\mathbf{r}_j}{dt} = \frac{\mathbf{p}_j}{m}$$

$$\frac{d\mathbf{p}_j}{dt} = \mathbf{F}_j$$



F: evaluate ½N(N−1) pairs

Important features of an integrator:

- minimal need to compute forces (a very expensive calculation)
- good stability for large time steps
- good accuracy
- conserves energy and momentum
- time-reversible
- symplectic: conserves volume in phase space

Molecular Dynamics

Integrating the equations of motion

Discretize in time (n steps), with time step Δt:

$$r_i(t_0) \rightarrow r_i(t_0+\Delta t) \rightarrow r_i(t_0+2\Delta t) \rightarrow \ldots \rightarrow r_i(t_0+n\Delta t)$$

Simplest solution is to use a Taylor expansion:

$$r_i(t_0+\Delta t) = r_i(t_0) + v_i(t_0)\Delta t + \frac{1}{2}a_i(t_0)\Delta t^2 + O(\Delta t^3)$$

$$v_i(t_0+\Delta t) = v_i(t_0) + a_i(t_0)\Delta t + O(\Delta t^2)$$

But it does not work too well → **Unstable and inaccurate!**

Molecular Dynamics

Verlet algorithm

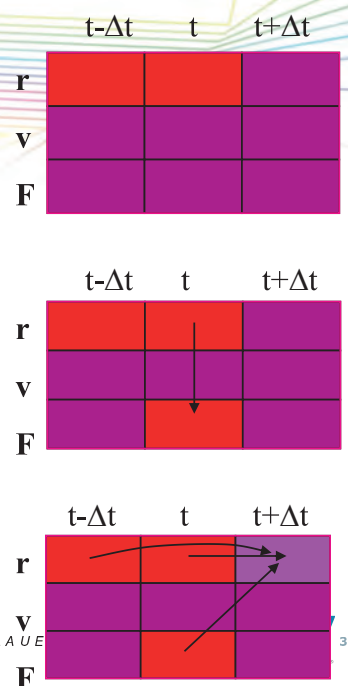
$$r_i(t_0+\Delta t) = r_i(t_0) + v_i(t_0)\Delta t + \frac{1}{2}a_i(t_0)\Delta t^2 + \frac{1}{3!}\dddot{r}_i(t_0)\Delta t^3 + O(\Delta t^4)$$

$$r_i(t_0-\Delta t) = r_i(t_0) - v_i(t_0)\Delta t + \frac{1}{2}a_i(t_0)\Delta t^2 - \frac{1}{3!}\dddot{r}_i(t_0)\Delta t^3 + O(\Delta t^4)$$

$$r_i(t_0+\Delta t) + r_i(t_0-\Delta t) = 2r_i(t_0) + a_i(t_0)\Delta t^2 + O(\Delta t^4)$$

Error is $O(\Delta t^4)$. This algorithm does not use the velocities, but they can be derived from:

$$r_i(t_0+\Delta t) - r_i(t_0-\Delta t) = 2v_i(t_0) + O(\Delta t^3)$$



Molecular Dynamics

Leap-frog algorithm

$$r_i(t_0+\Delta t) = r_i(t_0) + v_i(t_0+\Delta t/2)\Delta t$$

$$v_i(t_0+\Delta t/2) = v_i(t_0-\Delta t/2) + a_i\Delta t$$

- Equivalent to Verlet!
- Velocities not given at same time as positions, but can be recasted to give them at same time (velocity Verlet).
- Verlet and equivalent alternatives are simple, efficient, stable and reasonably accurate.
- Time-reversible and symplectic → low drifts in energy.

Molecular Dynamics

Other integration algorithms

Many numerical methods are available to integrate the equations of motion, but we can discard all those requiring more than one evaluation of the forces (too time consuming). Apart from Verlet and equivalent, the other type of algorithms that are used some times are the predictor-corrector:

- Predict values at next time-step:

$$\boldsymbol{r}^p(t+\Delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t)\Delta t + \frac{1}{2}\boldsymbol{a}(t)\Delta t^2 + \frac{1}{6}\dot{\boldsymbol{a}}(t)\Delta t^3 + \ldots; \quad \boldsymbol{a}^p(t+\Delta t) = \boldsymbol{a}(t) + \dot{\boldsymbol{a}}(t)\Delta t + \ldots$$

$$\boldsymbol{v}^p(t+\Delta t) = \boldsymbol{v}(t) + \boldsymbol{a}(t)\Delta t + \frac{1}{2}\dot{\boldsymbol{a}}(t)\Delta t^2 + \ldots \quad ; \quad \dot{\boldsymbol{a}}^p(t+\Delta t) = \dot{\boldsymbol{a}}(t) + \ldots$$

- Then use $r^p(t+\Delta t)$ to calculate correct accelerations, $a^c(t+\Delta t)$, and correct.

More accurate for small Δt, but not time reversible → larger drifts in energy.

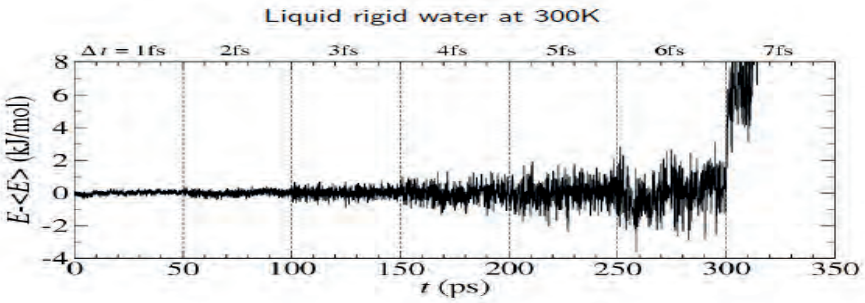
Molecular Dynamics

Time step

Too small: inefficient phase space sampling

Too large: numerical instabilities → integrator divergence → MD crash

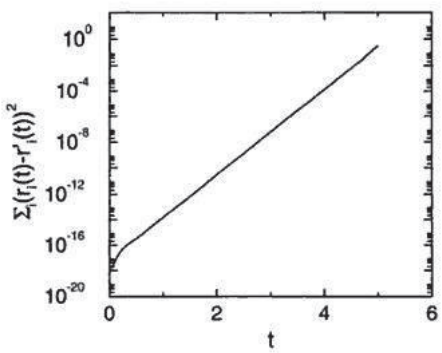
Hints: - Δt should not be larger than the meantime between collisions
- Flexible molecules and rigid bonds: 2 fs – Flexible bonds: 1 fs
- Total energy should be conserved → $\langle dE^2 \rangle^{1/2}/E < 10^{-4}$



Molecular Dynamics

Lyapunov instability

- Two initially close trajectories will diverge exponentially: impossible to calculate a true trajectory!
- Fortunately we don't want to predict a trajectory, but statistical properties.
- We assume that the results of our simulation are representative of a true trajectory in phase space, even if we cannot tell which.

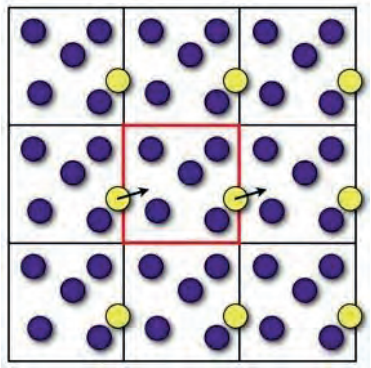


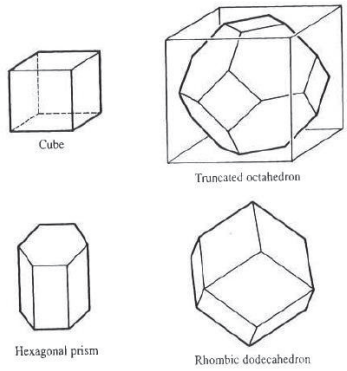
Difference between trajectories of two simulations of 1000 LJ atoms where the velocities of 2 atoms were changed by $+10^{-10}$ and -10^{-10} (r.u.) (from Frenkel & Smit, 2002)

Molecular Dynamics

Periodic Boundary Conditions

- Enables bulk properties to be computed using a limited number of atoms
- Remove surface effects ~ $N^{-1/3}$ (49% for 1000 atoms, 6% for 10^6)
- When an atom leaves the simulation cell, it is replaced by another with the same velocity, entering from the opposite cell face (N conserved)





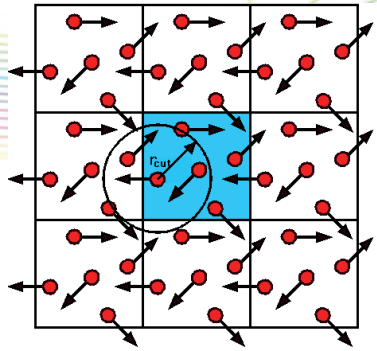
Molecular Dynamics

Periodic Boundary Conditions

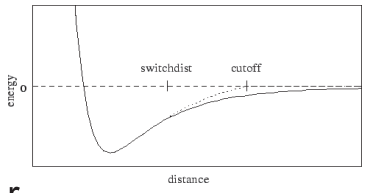
- Only allowed fluctuations are those whose wavelengths are compatible with the box length.
- Longest wavelength $\lambda = L$ → problem if long wavelength fluctuations are important (e.g. phase transitions).
- When simulating solids, the strain field generated by any inhomogeneity will be artificially truncated and modified by the boundary.
- A macromolecule may interact with its own image if the box is too small (e.g. recommendation of using a minimum layer of 10 Å of water when simulating DNA, Souza & Ornstein, *Biophys. J.* (1997)).
- In some cases spurious correlations when dealing with charged or strongly polar solutes.

Molecular Dynamics

Non-bonded interactions



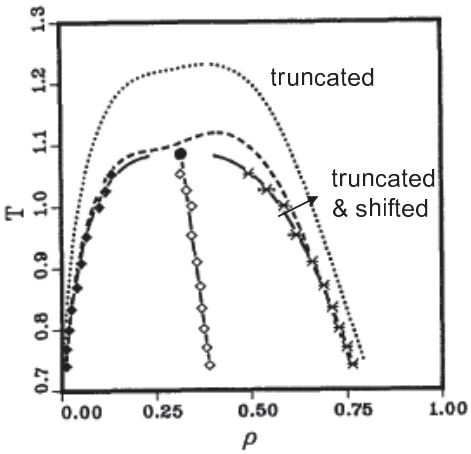
- Most time-consuming part of the simulation:
  - Bonded interactions → O(N)
  - Non-bonded interactions → $O(N^2)$
  - Not feasible to include interactions with all images
- Minimum image convention:
  - an atom just sees the closest image
- Use a non-bonded cutoff:
  - truncate the potential
  - consider interactions only inside a sphere of radius r_{cut}
  - discontinuity in force and energy calculation (truncated)



Molecular Dynamics

Non-bonded forces: Truncation effects

- No truncation method is ideal, but differences are less significant the longer R_c is chosen. So the cutoff should be as large as possible (Steinbach & Brooks, *J. Comput. Chem.* (1994)).
- Do not cut long-range interactions (electrostatics), even using large R_c (e.g. Yonetani, *J. Chem. Phys.* (2006) shows larger artifacts for a liquid water simulation with R_c=18Å than with R_c=9Å)
- Quite often it does not matter too much which truncation method is used, but take care with some particular conditions. E.g.: Close to critical points.
- Typically $R_c \approx 2.5\sigma_{LJ} \approx$ 10-12 Å



Phase diagram of LJ fluid (Smit, J. Chem. Phys. 1992)

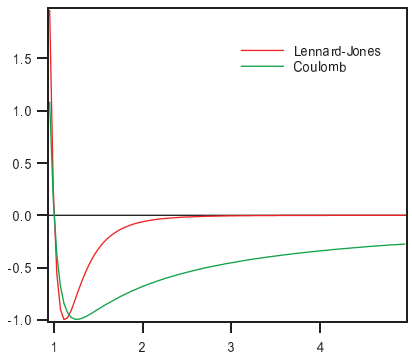
Molecular Dynamics

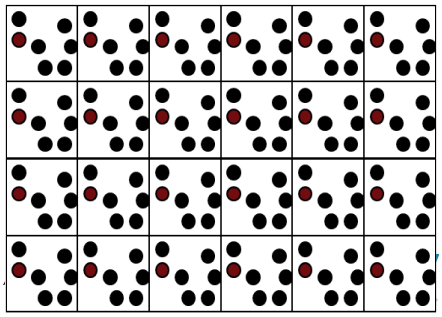
Electrostatic interactions

- Very long (beyond primary cell) range
- 1/r does not die off as quickly as volume grows

$$\int_{r_c}^{\infty} \frac{1}{r} 4\pi r^2 dr = \infty$$

- Finite only because + and – cancel
- Methods:
  - treat surroundings as dielectric continuum
  - full lattice sum → Ewald sum

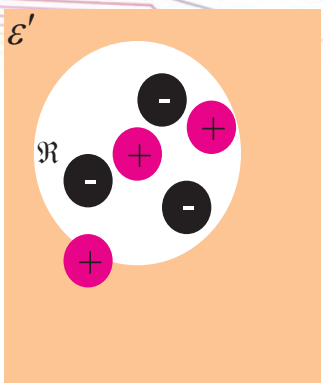




Molecular Dynamics

Electrostatics: Reaction Field

- Considers all the interactions up to a certain cutoff, R_c, and use a mean-field approximation beyond.
- For each particle calculate directly all interactions with the atoms inside the sphere and the total dipole moment of the sphere μ_s.
- Outside the sphere the medium is considered to be a continuum with a dielectric constant ε_s that produces a reaction field proportional to μ_s.
- The RF contribution to the energy will be:

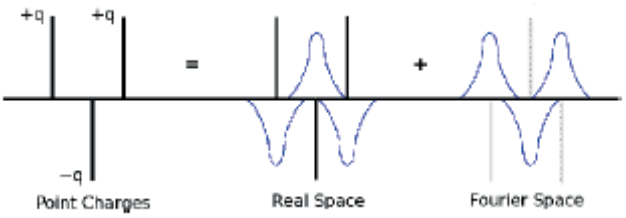


$$U_{rf} = \frac{1}{2}\sum_{i=1}^{N} \mu_i \cdot \left[\frac{2(\varepsilon'-1)}{2\varepsilon'+1} \frac{1}{r_c^3} \sum_{j \in \Re_i} \mu_j \right]$$

Molecular Dynamics

Electrostatics: Ewald sum

Reference method to treat long-range interactions, often in one of its particle-mesh variants: PME, SPME, or PPPE.



At each point charge add a gaussian distribution of opposite sign to screen it

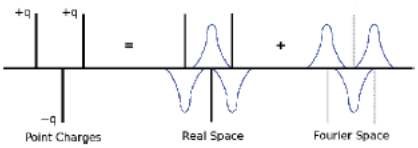
$$\rho = -q_i \left(\frac{\alpha}{\pi}\right)^{3/2} \exp\left(-\alpha r^2\right)$$

The electrostatic potential caused by the screened charges is now short-ranged and can be treated in real space:

$$U_{sr} = \frac{1}{2}\sum_{i \neq j}^{N} q_i q_j \operatorname{erfc}\left(\sqrt{\alpha} r_{ij}\right) / r_{ij}$$

Molecular Dynamics

Electrostatics: Ewald sum



The compensating Gaussian distribution can be treated in **k**-space

$$U_{lr} = \frac{1}{2V}\sum_{\boldsymbol{k} \neq 0} \frac{4\pi}{k^2} |\rho(\boldsymbol{k})|^2 \exp\left(-k^2/4\alpha\right)$$

$$\text{with } \rho(\boldsymbol{k}) = \sum_{i=1}^{N} q_i \exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i\right)$$

At this point we have also considered the unphysical interaction of the point charge with its compensating cloud, so we need to add a correction to this self –interaction:

$$U_{\text{self}} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \sum_{i=1}^{N} q_i^2$$

Molecular Dynamics

Thermodynamic ensembles

- The integration of the equations of motion keeps constant N, V, and E → **microcanonical** ensemble.
- But integration errors, force fluctuations and inconsistencies in the forces (e.g. generated by the cutoff) may cause slow drifts in the total energy.
- E=const, but not K and U, so systems not in equilibrium will go to equilibrium while the temperature changes.
- We may prefer to work at constant T or P or both to compare to experiment.

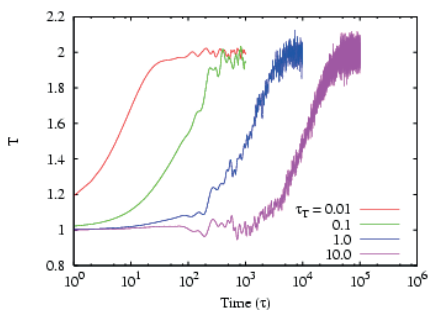
We can modify the Lagrangian or couple the system to a heat or pressure bath to keep T, P or both constant!

Molecular Dynamics

Thermodynamic ensembles

Several thermostats allow to do **NVT** simulations: velocity scaling, Berendsen, Andersen, Nosé-Hoover, but not all of them sample strictly the correct thermodynamic ensemble.

There are also different barostats to do **NPT** simulations.



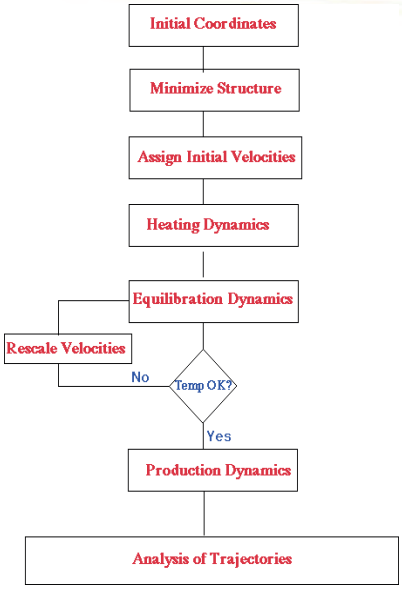
Berendsen: Simple, fast response and flexible, but no canonical ensemble.

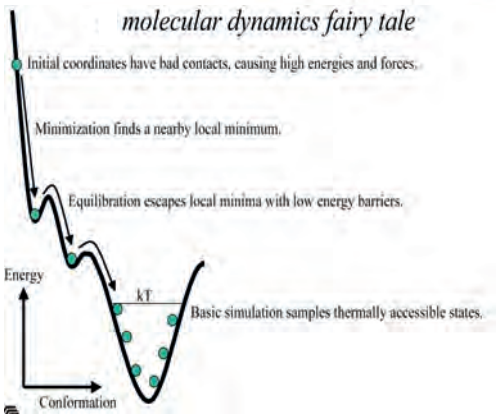
$$\boldsymbol{a} = \boldsymbol{F}/m + \frac{1}{2\tau_T}\left(\frac{T_B}{T(t)} - 1\right)\boldsymbol{v}$$

Nose-Hoover: Extended lagrangian. Sample correct canonical ensemble, but goes to T_{ref} in an oscillatory way (poorer T control).

Molecular Dynamics

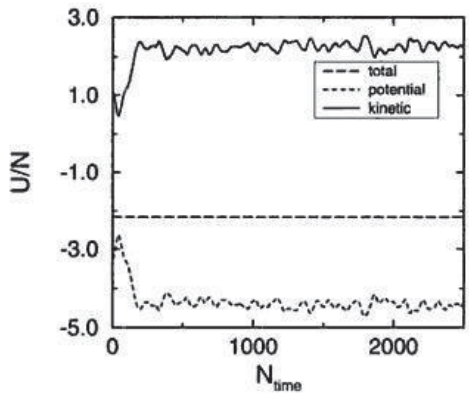
Protocol





Molecular Dynamics

Equilibration



CHECK:

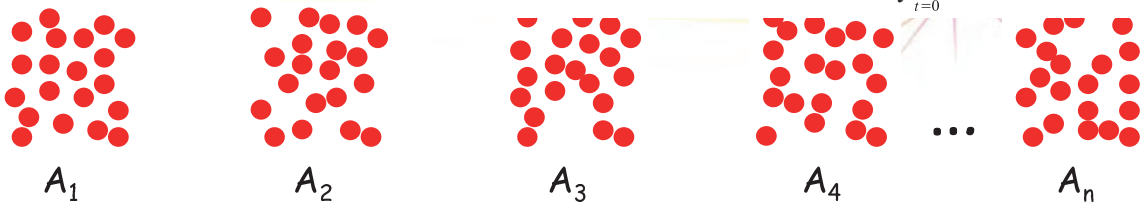
- Energy conservation
- Drift in temperature
- Equilibration:
  - exchange between K and U
  - K, U, ρ, structure converge

· **Production:** Simulation time long enough to sample the desired property, but **also to apply ergodic hypothesis**.

Molecular Dynamics

Ergodic principle

$$\langle A \rangle = \int\int dp^N dr^N A(p^N, r^N)\rho(p^N, r^N) = \lim_{\tau \to +\infty} \frac{1}{\tau} \int_{t=0}^{\tau} A\left(p^N(t), r^N(t)\right) dt$$



- Ensemble = collection of large number of replicas:
  - each replica has the same macroscopic parameters (e.g. NVT)
  - differ microscopically → fluctuation
  - ⟨ A ⟩ = average over all replicas → **ensemble average**
- Molecular dynamics:
  - each time step generates a new configuration of a **single** replica
  - ⟨ A ⟩ = average over all configurations → **time average**
  - **Ergodic principle → ensemble average = time average**

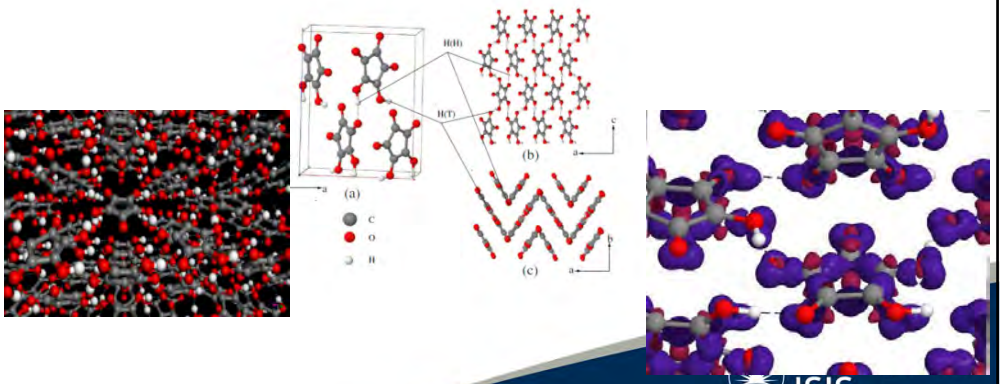
Thanks!

Questions?

Introduction to Electronic Structures

Sanghamitra Mukhopadhyay

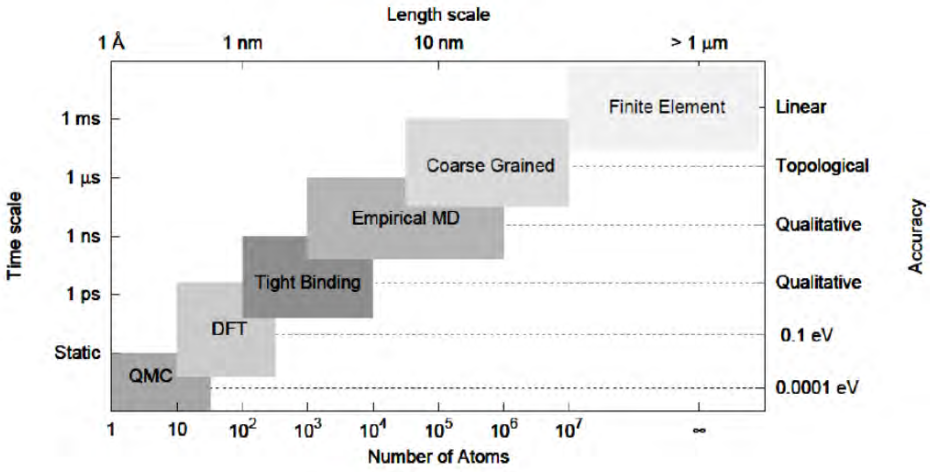
ISIS Facility, STFC Rutherford Appleton Laboratory
Oxfordshire



ISIS

Level of Approximations

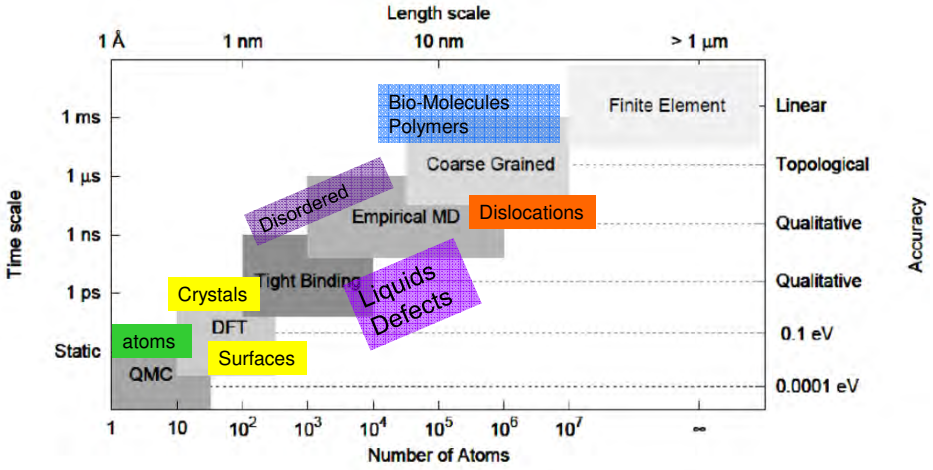
Comparison between computational costs, accuracy and generality



Multiscale simulations in Mat. Sc., J. Grotendorst et al., NIC Series, Vol 12 (2009).

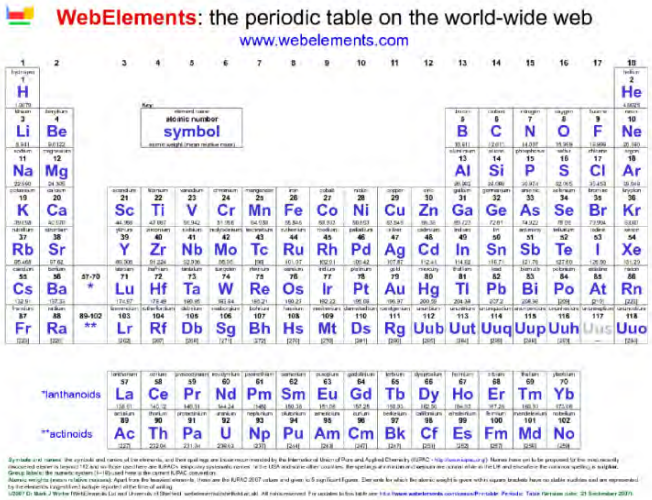
Level of Approximations

Comparison between computational costs, accuracy and generality



Multiscale simulations in Mat. Sc., J. Grotendorst et al., NIC Series, Vol 12 (2009).

Alphabets of the Materials World



Properties depend on electronic structures

Science & Technology Facilities Council
ISIS

Essential Ingredients of Electronic Structures

- Consider nucleus as point mass with positive charge
- Consider each electron as point mass with negative charge
- Apply Coulomb's interactions between nucleus-electron, electron-electron
- Apply force law, laws of motion, conservation of energy, momentum

$$\hat{H} = \overbrace{\hat{T}_n}^{\text{Nucl. kin. energy}} + \overbrace{\hat{T}_e}^{\text{Elec. kin. energy}} + \overbrace{\hat{v}_{ne}}^{\text{Elec.-nucl. attraction}} + \overbrace{\hat{v}_{ee}}^{\text{Elec.-elec repulsion}} + \overbrace{\hat{v}_{nn}}^{\text{Nucl.-nucl repulsion}}$$

Science & Technology Facilities Council ISIS

Essential Ingredients of Electronic Structures

- Consider nucleus as point mass with positive charge
- Consider each electron as point mass with negative charge
- Apply Coulomb's interactions between nucleus-electron, electron-electron
- Apply force law, laws of motion, conservation of energy, momentum

$$\hat{H} = -\sum_{a=1}^{Nb} \frac{\hbar^2}{2\,\mathbf{M_a}} \vec{\nabla}^2_{\vec{R}_a} - \sum_{i=1}^{N} \frac{\hbar^2}{2\,\mathbf{m_e}} \vec{\nabla}^2_{\vec{r}_i}$$
$$- \frac{e^2}{4\pi\epsilon_0} \sum_{a=1}^{Nb} \sum_{i=1}^{N} \frac{Z_a}{|\vec{R}_a - \vec{r}_i|} + \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^{N} \sum_{j=1}^{i-1} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{e^2}{4\pi\epsilon_0} \sum_{a=1}^{Nb} \sum_{b=1}^{a-1} \frac{Z_a Z_b}{|\vec{R}_a - \vec{R}_b|}$$

Science & Technology Facilities Council ISIS

Solving the Schrödinger equation for Many-body Electron System

Schrödinger equation:

$$\hat{H}\Psi = E\Psi$$

$$\frac{-\hbar^2}{2m_e}\Psi + \hat{V}\Psi = E\Psi$$

$$\begin{aligned}
\hat{H}_e &= \hat{T}_e + \hat{v}_{ne} + \hat{v}_{ee} + v_{nn} \\
&= \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m_e}\vec{\nabla}^2_{\vec{r}_i} - \frac{e^2}{4\pi\epsilon_0}\sum_{a=1}^{Nb}\frac{Z_a}{|\vec{R}_a - \vec{r}_i|}\right] + \sum_{i=1}^{N}\sum_{j=1}^{i-1}\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\vec{r}_i - \vec{r}_j|} + v_{nn} \\
&= \underbrace{\sum_{i=1}^{N}\hat{h}_1(\vec{r}_i)}_{\text{simple}} + \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{i-1}\hat{h}_2(\vec{r}_i,\vec{r}_j)}_{\text{difficult}} + v_{nn}
\end{aligned}$$

ISIS

Approximate Solution: DFT

Interacting electrons in an external potential

Non-interacting electrons in an effective potential

These non-interacting electrons are called as Kohn-Sham particles

Approximate e-e interactions with
- Local density approximation
- Generalized gradient approximation
- Hybrid …

Problem is exact form of this functionals are unknown.

Science & Technology Facilities Council ISIS

Kohn-Sham Equation

$$\{\frac{1}{2}\nabla^2+V_{\text{ext}}(\vec{r})+V_H(\vec{r})+V_{\text{xc}}([n(\vec{r})])\}\varphi_i(\vec{r})=\epsilon_i\varphi_i(\vec{r})$$

Electron density $\overline{n(\vec{r})}=\sum_i^{occ}|\varphi_i^2(\vec{r})|$

Hartree energy is the self interaction energy of the density n(r) treated as a classical charge density

Hartree exchange can be calculated exactly, but e-e correlation is unknown

Science & Technology Facilities Council ISIS

Local Density Approximation

The $\hat{v}_{\text{xc}}[n]$ functionals

- **LDA/LSDA** Local Density Approximation
  $\hat{v}_{\text{xc}}[n]$ = homogeneous ē gas
  Perdew et Zunger (PZ), Vosko Wilkes and Nusiar (VWN)
  - Easy, fast
  - OK for Fermi ē of metals
  - Bad for core ē
  - Only local $f(n(\vec{r}))$
    exch.-corr. : non local $f(\vec{r},\vec{r}')$
  - Self-inter. pb
  - Overdelocalize
  - Underestimates gaps
  - No dispersion forces

Science & Technology Facilities Council ISIS

Generalised Gradient Approximation

The $\hat{v}_{xc}[n]$ functionals

- **GGA** Generalized Gradient Approximation
  Add semi-local terms : $f(n(\vec{r}), \vec{\nabla}_{\vec{r}}(n(\vec{r}))$
  Perdew-Burke-Ernzerhof (PBE, PBESOL), Perdew-Wang 91 (PW91), BLYP, second order (SOGGA), Perdew-Wang (PW), Wu-Cohen (WC)
  - A little better than LDA

Science & Technology Facilities Council ISIS

Hybrid Functionals

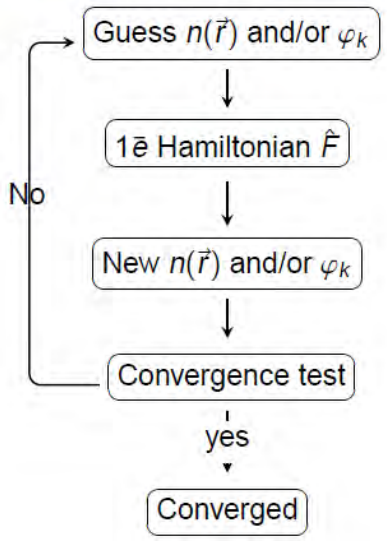
The $\hat{v}_{xc}[n]$ functionals

- **Hybrids functionals**
  Mix of LDA - GGA - HF exact exchange
  B3LYP, B3PW, PBE0 and PBESOL0, B1PW, B1WC
  - Good structures
  - Part self-inter. pb
  - Gaps
  - No dispersion forces
  - More costly
  - Cooking recipe

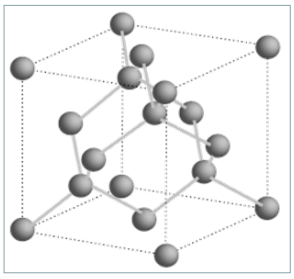
Science & Technology Facilities Council ISIS

Solving Electronic Structures: Recipe

1ē Hamiltonian $\hat{F}$ depends on its solution ⇒ SCF



Periodic Solid



Crystal structure = lattice +basis

Geometry of lattice can be obtained from X-ray or neutron diffraction experiments

Database of CIF files:

National Chemical Database Service

http://cds.rsc.org/

Inorganic Crystal Structure Database

https://www.fiz-karlsruhe.de/de/leistungen/kristallographie/icsd.html

Basis sets

One electron projected on finite basis set

- Plane wave (CASTEP, VASP)
- Linear Combination of Atomic Orbitals (LCAO) (CRYSTAL, GAUSSIAN)
- Atom centered numerical basis set (SIESTA)

Treatment of inner electrons

- Pseudopotentials
- All electrons
- Atomic functions with muffin-tin spheres

Science & Technology Facilities Council ISIS

Things to Remember

Test numerical parameters

- Type of pseudopotential
- Heavy elements : relativistic corrections
- Plane waves : energy cutoff
- Atomic basis sets : basis set quality
- $\vec{k}$ points convergence
- Metals : smearing temperature
- Phonons : geom. opt. convergence (negative phonons)

Science & Technology Facilities Council ISIS

Things to Remember

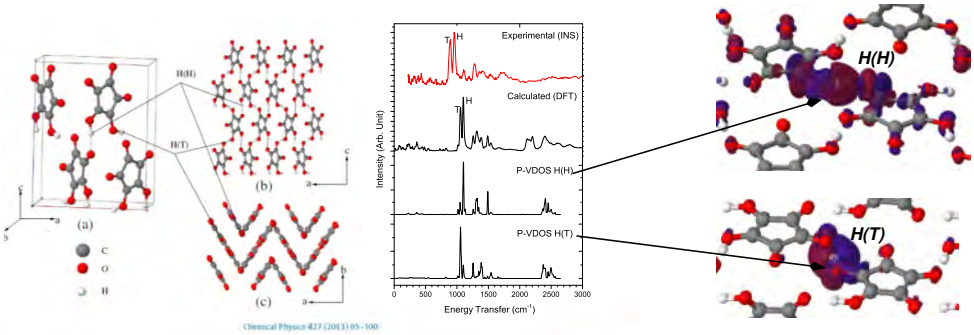
Do not use as black box

Always test for validity and computational parameters

Science & Technology Facilities Council
ISIS

First Principles Simulations : Example

Chemical insight from DFT based calculations



Experimental INS compared with Calculations

Response charge densities of hydrogen ions

Contents lists available at
Chemical Physics
journal homepage: www.elsevier.com/locate/chemphys
ELSEVIER

Ferroelectric behaviour in solid croconic acid using neutron scattering and first-principles density functional theory

S. Mukhopadhyay [a,d,e], M.J. Gutmann [a], M. Jura [a], D.B. Jochym [b], M. Jimenez-Ruiz [c], S. Sturniolo [b], K. Refson [b], F. Fernandez-Alonso [a,e]

[a] ISIS Facility, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom
[b] Scientific Computing Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom
[c] Institut Laue Langevin, 6 rue Jules Horowitz 38042, Grenoble Cedex 9, France
[d] Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom
[e] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

Science & Technology Facilities Council
ISIS

First Principles Simulations : Example

Diffraction from powdered sample compared with DFT

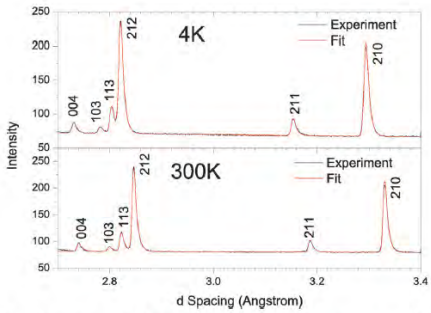


Fig. 2 ND patterns (black) and associated fits (red) obtained on IRIS at two selected temperatures. The numbers on the figure denote (*hkl*) reflections.

	Temperatures	
	10 K	
Parameters	ND	Calculated
O–H(H)	1.008 (±1.4)	1.040 (3.2%)
O–H(T)	1.005 (±2.4)	1.044 (3.9%)
O···H(H)	1.604 (±1.4)	1.476 (*8.0*%)
O···H(T)	1.611 (±1.4)	1.531 (*5.0*%)
O···O(H)	2.594 (±1.1)	2.517 (*3.0*%)
O···O(T)	2.612 (±1.1)	2.553 (*2.2*%)

PCCP

PAPER

View Article Online

CrossMark

Cite this: *Phys. Chem. Chem. Phys.*, 2014, 16, 26234

Hydrogen-bond structure and anharmonicity in croconic acid

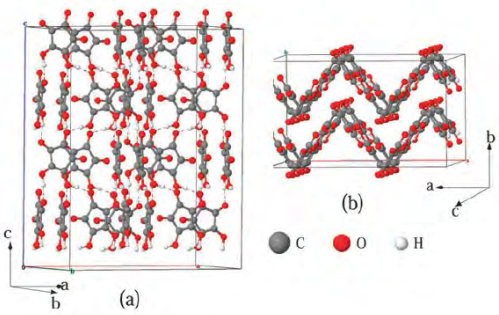
Sanghamitra Mukhopadhyay,*[ab] Matthias Gutmann[a] and Felix Fernandez-Alonso[bc]

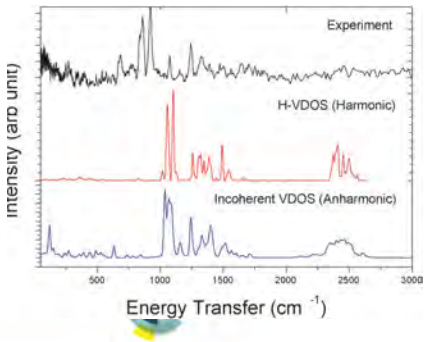
ence & Technology Facilities Council

SIS

First Principles Simulations : Example

Anharmonic DOS is compared with INS experiments





PCCP

PAPER

View Article Online

CrossMark

Cite this: *Phys. Chem. Chem. Phys.*, 2014, 16, 26234

Hydrogen-bond structure and anharmonicity in croconic acid

Sanghamitra Mukhopadhyay,*[ab] Matthias Gutmann[a] and Felix Fernandez-Alonso[bc]

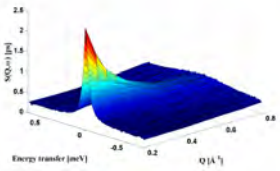
ence & Technology Facilities Council

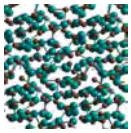
SIS

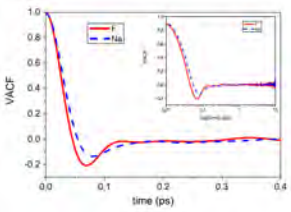
Computational Modelling: Molten Salt

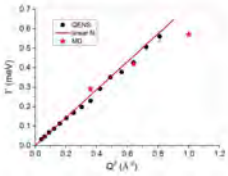
Single particle dynamics from molten salt

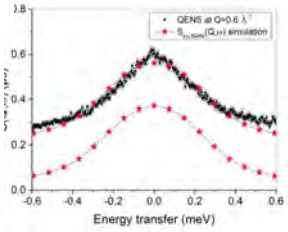
OSIRIS Experiment is compared with DFT-MD

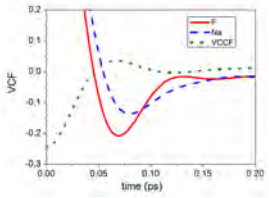












THE JOURNAL OF CHEMICAL PHYSICS **144**, 014503 (2016)

Quasielastic neutron scattering measurements and *ab initio* MD-simulations on single ion motions in molten NaF

F. Demmel[1] and S. Mukhopadhyay[1,2]
[1]*ISIS Facility, Rutherford Appleton Laboratory, Didcot OX11 0QX, United Kingdom*
[2]*Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom*

y Facilities Council

Acknowledgements

- Felix Fernandez-Alonso — ISIS, RAL, UK
- Matthias Gutmann — ISIS, RAL, UK
- Keith Refson — CSED, RAL, UK
- Dominik Jochym — CSED, RAL, UK
- Jeff Armstrong — ISIS, RAL, UK

Science & Technology Facilities Council
ISIS

Introduction to Lattice Dynamics

Keith Refson
Royal Holloway, University of London

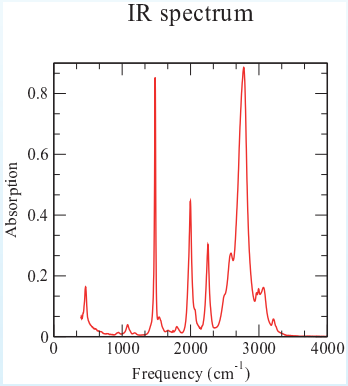
November 2016

Motivations for *ab initio* lattice dynamics I

● Motivations for *ab initio* lattice dynamics I
● Motivations for *ab initio* lattice dynamics II
Lattice Dynamics of Crystals
ab initio Lattice Dynamics
Examples
Modelling of spectra

Motivations from experimental spectroscopy:

- Vibrational spectroscopy is sensitive probe of structure and dynamics of materials.
- All experimental methods (IR, Raman, INS, IXS) provide *incomplete* information.
  - IR and Raman have inactive modes
  - Hard to distinguish fundamental and overtone (multi-phonon) processes in spectra
  - No experimental technique provides complete eigenvector information $\Rightarrow$ mode assignment based on similar materials, chemical intuition, guesswork.
- Hard to find accurate model potentials to describe many systems
- Fitted force-constant models only feasible for small, high symmetry systems.



● Motivations for *ab initio* lattice dynamics I
● Motivations for *ab initio* lattice dynamics II
Lattice Dynamics of Crystals
ab initio Lattice Dynamics
Examples
Modelling of spectra

Motivations for *ab initio* lattice dynamics II

Motivations from predictive modelling

- Lattice dynamics (LD) calculation is direct test of stability or otherwise of putative structure.
- Thermodynamics: LD can compute zero-point energy (ZPE) and phonon entropy contributions to Free energy
- LD gives direct information on interatomic forces.
- Some phase transitions are direct result of “soft” phonon mode instability.
- Electron - phonon coupling is origin of (BCS) superconductivity.

● Motivations for *ab initio* lattice dynamics I
● Motivations for *ab initio* lattice dynamics II
Lattice Dynamics of Crystals
● References
● Monatomic Crystal in 1d (I)
● Monatomic Crystal (II)
● Diatomic Crystal - Optic modes
● Characterization of Vibrations in Crystals
● Dispersion, DOS and Gamma
● Formal Theory of Lattice Dynamics
● The Force Constant Matrix
● The Dynamical Matrix
● Formal Theory of Lattice Dynamics II
● Quantum Theory of Lattice Modes
ab initio Lattice Dynamics
Examples
Modelling of spectra

Lattice Dynamics of Crystals

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

Examples

Modelling of spectra

References

Books on Lattice Dynamics

- M. T. Dove *Introduction to Lattice Dynamics*, CUP. - elementary introduction.
- J. C. Decius and R. M. Hexter *Molecular Vibrations in Crystals* - Lattice dynamics from a spectroscopic perspective.
- Horton, G. K. and Maradudin A. A. *Dynamical properties of solids* (North Holland, 1974) A comprehensive 7-volume series - more than you'll need to know.
- Born, M and Huang, K *Dynamical Theory of Crystal Lattices*, (OUP, 1954) - The classic reference, but a little dated in its approach.

References on *ab initio* lattice dynamics

- K. Refson, P. R. Tulip and S. J Clark, Phys. Rev B. **73**, 155114 (2006)
- S. Baroni *et al* (2001), Rev. Mod. Phys **73**, 515-561.
- Variational DFPT (X. Gonze (1997) PRB **55** 10377-10354).
- Richard M. Martin *Electronic Structure: Basic Theory and Practical Methods: Basic Theory and Practical Density Functional Approaches Vol 1* Cambridge University Press, ISBN: 0521782856

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

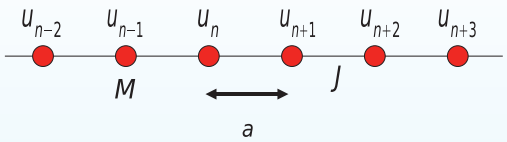
Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

Examples

Modelling of spectra

Monatomic Crystal in 1d (I)



Born-Von Karman boundary conditions

If only nearest neighbours interact:

$$E^{harm} = \frac{1}{2} J \sum_n (u_n - u_{n+1})^2$$

where $J = \Phi_{nn'}$ force constant

Equation of motion of n-th atom:

$$M \frac{\partial^2 u_n}{\partial t^2} = -\frac{\partial E^{harm}}{\partial u_n} = -J(2u_n - u_{n+1} - u_{n-1})$$

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

Examples

Modelling of spectra

Monatomic Crystal (II)

Solution is known to be the equation of a travelling wave

$$u_n(t)=\sum_k \tilde{u}_k \exp[i(kx-\omega_k t)] \qquad \forall k \in \mathrm{BZ}$$

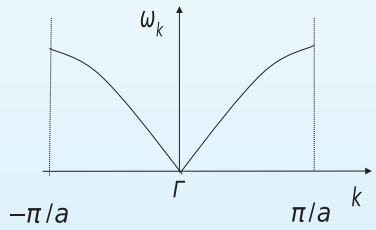
When this expression is inserted in the equation of motion one obtains:

$$M\omega_k^2 = 2J[1-\cos(ka)]$$

which leads to the following dispersion curve:

$$\omega_k = \sqrt{\frac{4J}{M}}|\sin(ka/2)|$$

$$\forall k \in \mathrm{BZ}$$



- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

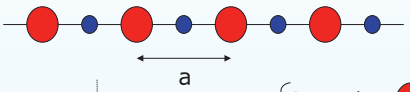
ab initio Lattice Dynamics

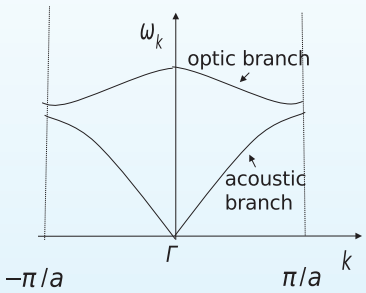
Examples

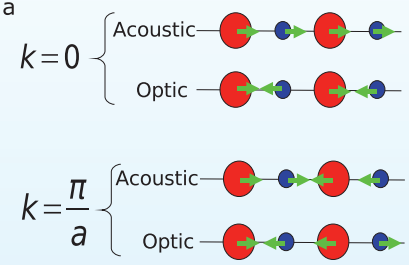
Modelling of spectra

Diatomic Crystal - Optic modes

More than one atom per unit cell gives rise to *optic modes* with different characteristic dispersion.







- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

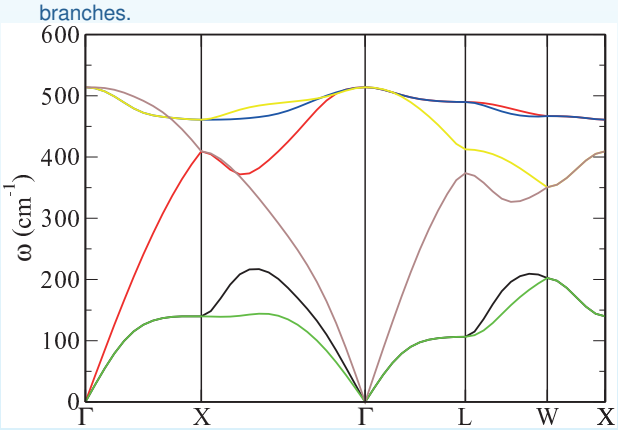
ab initio Lattice Dynamics

Examples

Modelling of spectra

Characterization of Vibrations in Crystals

- Vibrational modes in solids take form of *waves* with wavevector-dependent frequencies (just like electronic energy levels).
- $\omega(\boldsymbol{q})$ relations known as *dispersion curves*
- N atoms in prim. cell $\Rightarrow 3N$ *branches*.
- 3 *acoustic* branches corresponding to sound propagation as $\boldsymbol{q} \to 0$ and $3N-3$ *optic* branches.



- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

Examples

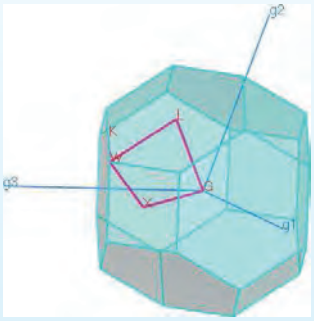
Modelling of spectra

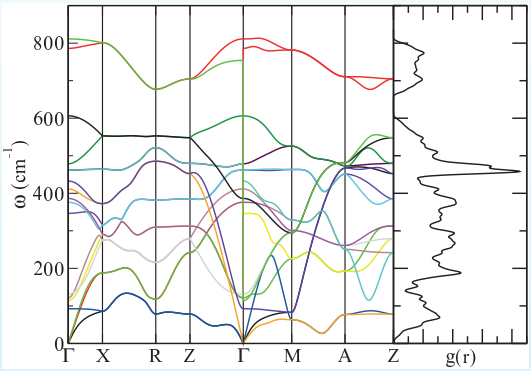
Dispersion, DOS and Gamma

Phonons represented across the Brillouin Zone by

Dispersion Curve $\omega_m(\boldsymbol{k})$ on high-symmetry lines

Phonon Density of States $g(\omega) = \int d\boldsymbol{k} \sum_m \delta(\omega_m(\boldsymbol{k}) - \omega)$





- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

Examples

Modelling of spectra

Formal Theory of Lattice Dynamics

- Based on expansion of total energy about structural equilibrium co-ordinates

$$E = E_0 + \sum_{\kappa,\alpha,a} \frac{\partial E}{\partial \boldsymbol{u}_{\kappa,\alpha,a}} . \boldsymbol{u}_{\kappa,\alpha,a} + \frac{1}{2} \sum_{\kappa,\alpha,\kappa',\alpha',a} \boldsymbol{u}_{\kappa,\alpha,a} . \Phi^{\kappa,\kappa'}_{\alpha,\alpha'} . \boldsymbol{u}_{\kappa',\alpha',a} + \ldots$$

  where $\boldsymbol{u}_{\kappa,\alpha,a}$ is the vector of atomic displacements from equilibrium.
- Notation: κ, κ' label atoms within a unit cell, a runs over all unit cells, α is Cartesian XYZ.
- $\Phi^{\kappa,\kappa'}_{\alpha,\alpha'}(a)$ is the matrix of *force constants*

$$\Phi^{\kappa,\kappa'}_{\alpha,\alpha'}(a) = \frac{\partial^2 E}{\partial \boldsymbol{u}_{\kappa,\alpha} \partial \boldsymbol{u}_{\kappa',\alpha'}}$$

- At equilibrium the forces $F_{\kappa,\alpha} = -\frac{\partial E}{\partial \boldsymbol{u}_{\kappa,\alpha}}$ are all zero so 1st term vanishes.
- In the *Harmonic Approximation* the 3rd and higher order terms are assumed to be negligible

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

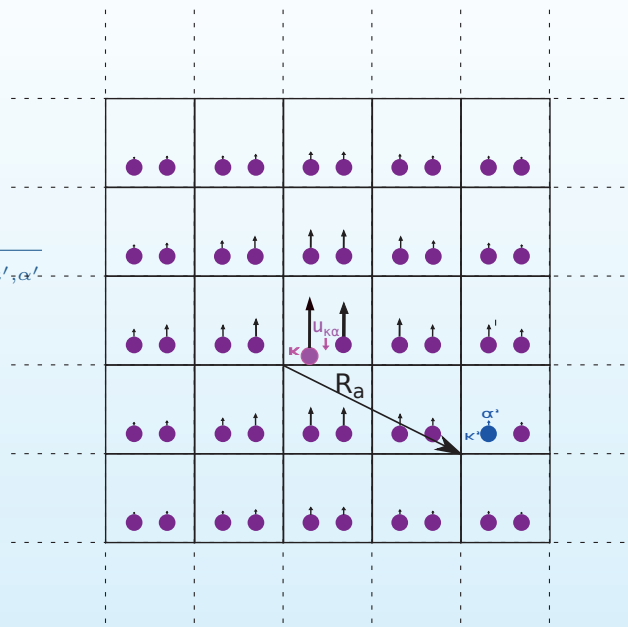
Examples

Modelling of spectra

The Force Constant Matrix

$$\Phi^{\kappa,\kappa'}_{\alpha,\alpha'}(a) = \frac{\partial^2 E}{\partial \boldsymbol{u}_{\kappa,\alpha} \partial \boldsymbol{u}_{\kappa',\alpha'}}$$

$$= -\frac{\partial F_{\boldsymbol{u}_{\kappa,\alpha,a}}}{\partial \boldsymbol{u}_{\kappa',\alpha',a}}$$



- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

Examples

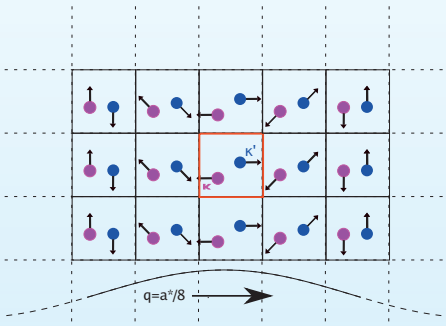
Modelling of spectra

The Dynamical Matrix

Assume *Born von Karman* periodic boundary conditions and substituting *plane-wave* $\boldsymbol{u}_{\kappa,\alpha} = \boldsymbol{\varepsilon}_{m\kappa,\alpha\boldsymbol{q}} \exp(i\boldsymbol{q}.\boldsymbol{R}_{\kappa,\alpha} - \omega t)$ yields eigenvalue equation:

$$D^{\kappa,\kappa'}_{\alpha,\alpha'}(\boldsymbol{q})\boldsymbol{\varepsilon}_{m\kappa,\alpha\boldsymbol{q}} = \omega^2_{m,\boldsymbol{q}}\boldsymbol{\varepsilon}_{m\kappa,\alpha\boldsymbol{q}}$$

$$D^{\kappa,\kappa'}_{\alpha,\alpha'}(\boldsymbol{q}) = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}} C^{\kappa,\kappa'}_{\alpha,\alpha'}(\boldsymbol{q}) = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}} \sum_a \Phi^{\kappa,\kappa'}_{\alpha,\alpha'}(a) e^{-i\boldsymbol{q}.\boldsymbol{R}_a}$$



- The **dynamical matrix** $D^{\kappa,\kappa'}_{\alpha,\alpha'}(\boldsymbol{q})$ is the Fourier transform of the force constant matrix.
- The solutions of the eigenvalue equation correspond to vibrational **modes**
- Mode frequency is square root of corresponding eigenvalue $\omega_{m,\boldsymbol{q}}$.

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals
- References
- Monatomic Crystal in 1d (I)
- Monatomic Crystal (II)
- Diatomic Crystal - Optic modes
- Characterization of Vibrations in Crystals
- Dispersion, DOS and Gamma
- Formal Theory of Lattice Dynamics
- The Force Constant Matrix
- The Dynamical Matrix
- Formal Theory of Lattice Dynamics II
- Quantum Theory of Lattice Modes

ab initio Lattice Dynamics

Examples

Modelling of spectra

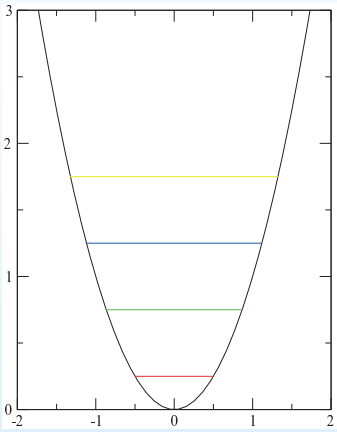
Formal Theory of Lattice Dynamics II

- The dynamical matrix is a $3N \times 3N$ matrix at each wavevector $\boldsymbol{q}$.
- $D^{\kappa,\kappa'}_{\alpha,\alpha'}(\boldsymbol{q})$ is a hermitian matrix $\Rightarrow$ eigenvalues $\omega^2_{m,\boldsymbol{q}}$ are real.
- $3N$ eigenvalues $\Rightarrow$ modes at each $\boldsymbol{q}$ leading to $3N$ branches in dispersion curve.
- The mode eigenvector $\boldsymbol{\varepsilon}_{m\kappa,\alpha}$ gives the atomic displacements, and its symmetry can be characterised by group theory.
- Given a force constant matrix $\Phi^{\kappa,\kappa'}_{\alpha,\alpha'}(a)$ we have a procedure for obtaining mode frequencies and eigenvectors over entire BZ.
- In 1970s force constants fitted to experiment using simple models.
- 1980s - force constants calculated from empirical potential interaction models (now available in codes such as GULP)
- mid-1990s - development of *ab initio* electronic structure methods made possible calculation of force constants with no arbitrary parameters.

● Motivations for *ab initio* lattice dynamics I
● Motivations for *ab initio* lattice dynamics II
Lattice Dynamics of Crystals
● References
● Monatomic Crystal in 1d (I)
● Monatomic Crystal (II)
● Diatomic Crystal - Optic modes
● Characterization of Vibrations in Crystals
● Dispersion, DOS and Gamma
● Formal Theory of Lattice Dynamics
● The Force Constant Matrix
● The Dynamical Matrix
● Formal Theory of Lattice Dynamics II
● Quantum Theory of Lattice Modes
ab initio Lattice Dynamics
Examples
Modelling of spectra

Quantum Theory of Lattice Modes

- The classical energy expression can be transformed into a quantum-mechanical Hamiltonian for nuclei.
- In harmonic approximation nuclear wavefunction is *separable* into product by mode transformation.
- Each *mode* satisfies harmonic oscillator Schroedinger eqn with energy levels $E_{m,n} = \left(n + \frac{1}{2}\right)\hbar\omega_m$ for mode m.
- Quantum excitations of modes known as *phonons* in crystal
- Transitions between levels n_1 and n_2 interact with photons of energy $(n_2 - n_1)\,\hbar\omega_m$, i.e. multiples of *fundamental* frequency ω_m.
- In *anharmonic* case where 3rd-order term not negligible, *overtone* frequencies are not multiples of fundamental.



● Motivations for *ab initio* lattice dynamics I
● Motivations for *ab initio* lattice dynamics II
Lattice Dynamics of Crystals
ab initio Lattice Dynamics
● Phonons from first principles
Examples
Modelling of spectra

ab initio Lattice Dynamics

● Motivations for *ab initio* lattice dynamics I
● Motivations for *ab initio* lattice dynamics II
Lattice Dynamics of Crystals
ab initio Lattice Dynamics
● Phonons from first principles
Examples
Modelling of spectra

Phonons from first principles

Methods for calculating phonons from ab initio calculations:

The Finite-Displacement method Displace each atom in turn;
Calculate force constants by numerical differentiation of atomic forces.
Restricted to commensurate (Γ-point) modes.

The Supercell method Extension of finite-displacement method, applied to large supercell.
Can compute phonons across entire BZ - full dispersion and DOS.

Density-Functional Perturbation Theory Compute force constants "directly" as derivatives.
Available at any q in BZ - full dispersion and DOS.

● Motivations for *ab initio* lattice dynamics I
● Motivations for *ab initio* lattice dynamics II
Lattice Dynamics of Crystals
ab initio Lattice Dynamics
Examples
● DFPT with interpolation - α-quartz
● Dispersion and DOS - Rutile TiO2
● DFPT vs Supercell method - bcc Tantalum
● Dynamical charges in NaHF2
Modelling of spectra

Examples

DFPT with interpolation - α-quartz

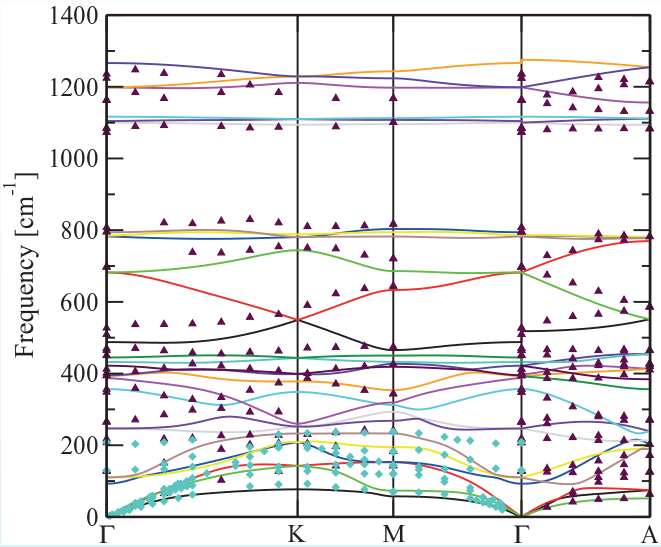
- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples
- DFPT with interpolation - α-quartz
- Dispersion and DOS - Rutile TiO2
- DFPT vs Supercell method - bcc Tantalum
- Dynamical charges in NaHF2

Modelling of spectra



Dispersion and DOS - Rutile TiO2

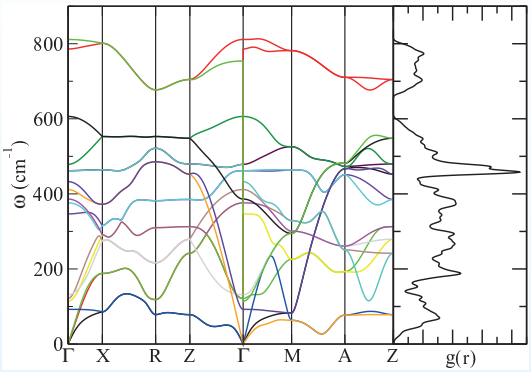
- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

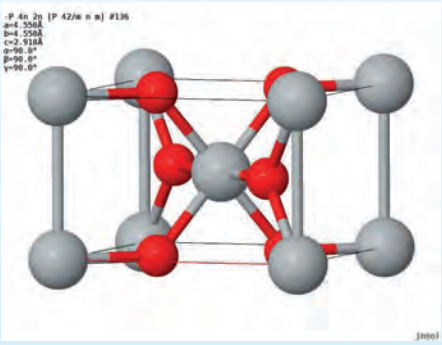
Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples
- DFPT with interpolation - α-quartz
- Dispersion and DOS - Rutile TiO2
- DFPT vs Supercell method - bcc Tantalum
- Dynamical charges in NaHF2

Modelling of spectra





- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

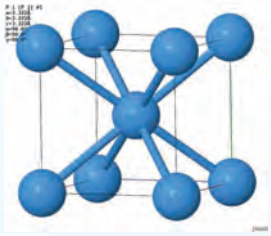
Lattice Dynamics of Crystals

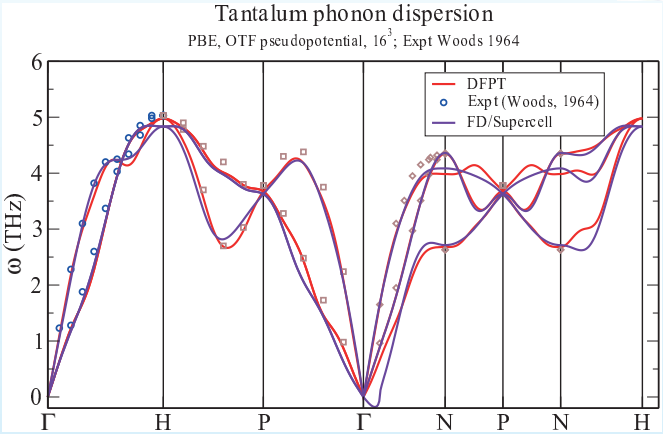
ab initio Lattice Dynamics

Examples
- DFPT with interpolation - α-quartz
- Dispersion and DOS - Rutile TiO2
- DFPT vs Supercell method - bcc Tantalum
- Dynamical charges in NaHF2

Modelling of spectra

DFPT vs Supercell method - bcc Tantalum





- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples
- DFPT with interpolation - α-quartz
- Dispersion and DOS - Rutile TiO2
- DFPT vs Supercell method - bcc Tantalum
- Dynamical charges in NaHF2

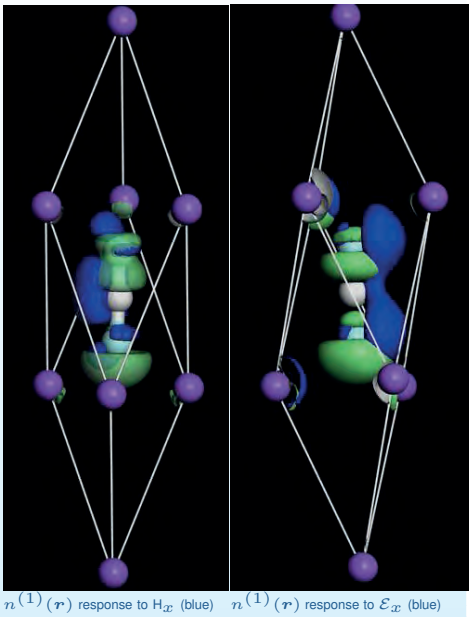
Modelling of spectra

Dynamical charges in NaHF2

- highly anisotropic Z^*

xy	z	
0.34	2.02	(H)
−0.69	−1.52	(F)
1.06	1.04	(Na)

- charge density $n^{(1)}(\boldsymbol{r})$ from DFPT shows electronic response under perturbation
- Charge on F ions moves in response to H displacement in z direction.



$n^{(1)}(\boldsymbol{r})$ response to H$_x$ (blue) and H$_z$ (green)

$n^{(1)}(\boldsymbol{r})$ response to $\mathcal{E}_x$ (blue) and $\mathcal{E}_z$ (green)

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples

Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

Modelling of spectra

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples

Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

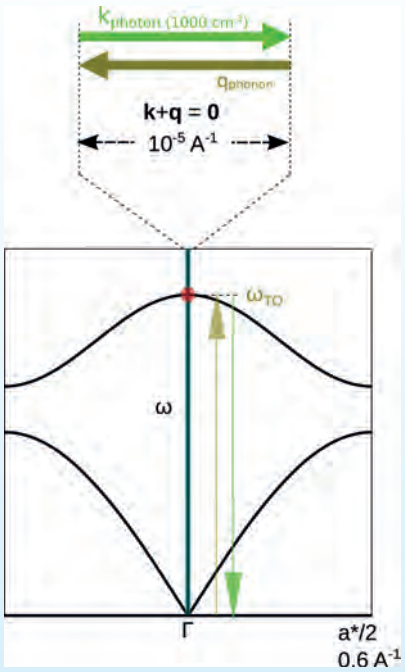
Interaction of photons with phonons

Energy conservation

- $E_s = h\nu_s = E_i + \delta E_{\text{phonon}} = h\nu_i + h\nu_{\text{phonon}}$

Momentum conservation

- $\boldsymbol{k}_s = \boldsymbol{k}_i + \delta \boldsymbol{k}_{\text{phonon}}$
- Typical lattice parameter $a_0 \approx 5.4$Å $\Rightarrow 0 \leq k_{\text{phonon}} < 0.6\ 1/$Å
- For a 10μm photon, $k = 10^{-5}\ 1/$Å
- $\Longrightarrow$ infra-red light interacts with phonons only very close to 0 (the "Γ" point).



- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples

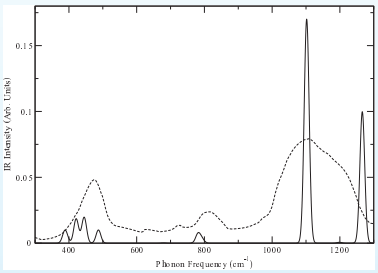
Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

Powder infra-red

In a powder infra-red absorption experiment, E field of light couples to *electric dipole* created by displacement of ions by modes. Depends on **Born effective charges** and **atomic displacements**.

Infra-red intensity $I_m = \left| \sum_{\kappa,b} \frac{1}{\sqrt{M_\kappa}} Z^*_{\kappa,a,b} u_{m,\kappa,b} \right|^2$

α-quartz



- Straightforward to compute peak *areas*.
- Peak shape modelling depends on sample and experimental variables.
- Multi-phonon and overtone terms less straightforward.

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples

Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

Raman Spectroscopy

Non-resonant Raman scattering depends on Raman activity tensor

$$I^{\text{Raman}}_{\alpha\beta} = \frac{d^3E}{d\varepsilon_\alpha d\varepsilon_\beta dQ_m} = \frac{d\epsilon_{\alpha\beta}}{dQ_m}$$

i.e. the activity of a mode is the derivative of the dielectric permittivity with respect to the displacement along the mode eigenvector.
CASTEP evaluates the Raman tensors using hybrid DFPT/finite displacement approach.
Raman calculation is fairly expensive $\Rightarrow$ and is not activated by default (though group theory prediction of active modes is still performed)
Parameter `calculate_Raman = true` in a `task=phonon` calculation.
Spectral modelling of IR spectrum is relatively simple function of activity.

$$\frac{d\sigma}{d\Omega} = \frac{(2\pi\nu)^4}{c^4} \left|\boldsymbol{e}_S.I.\boldsymbol{e}_L\right|^2 \frac{h(n_m+1)}{4\pi\omega_m}$$

with the thermal population factor

$$n_m = \left[exp\left(\frac{\hbar\omega_m}{k_B T}\right) - 1 \right]^{-1}$$

which is implemented in dos.pl “Raman calculation.

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

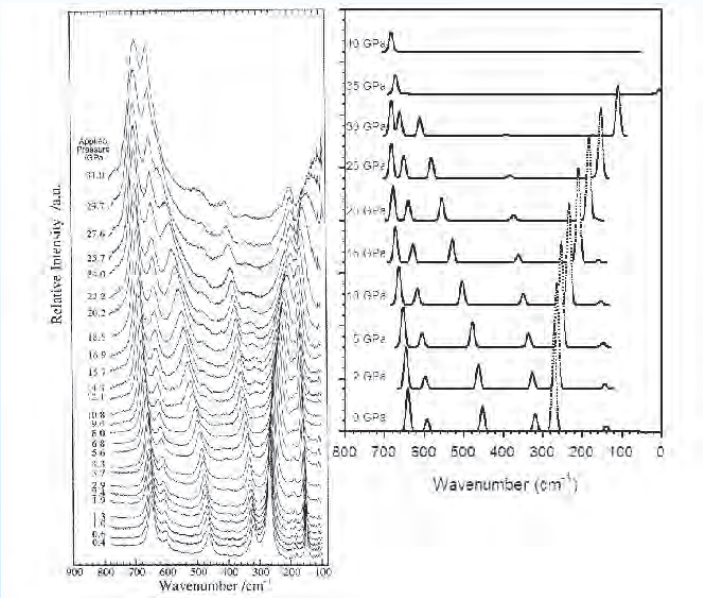
Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples

Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

Raman scattering of t-ZrO2



- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples

Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

Inelastic Neutron Scattering

Thermal neutrons have both energies and momenta of same magnitude as phonons (unlike photons). $\Rightarrow$ Can interact with phonons at any $\boldsymbol{q}$.
To model spectra need to treat scattering dynamics of incident and emergent radiation.
In case of INS interaction is between point neutron and nucleus - scalar quantity b depends only on nucleus – specific properties.

$$\frac{d^2\sigma}{dEd\Omega} = \frac{k_f}{k_i} b^2 S(\boldsymbol{Q}, \omega)$$

$\boldsymbol{Q}$ is scattering vector and ω is frequency - interact with phonons at same wavevector and frequency.
Full measured spectrum includes overtones and combinations and instrumental geometry and BZ sampling factors.
Need specific spectral modelling software to incorporate effects as post-processing step following CASTEP phonon DOS calculation.
A-Climax : A. J. Ramirez-Cuesta Comput. Phys. Comm. **157** 226 (2004))

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

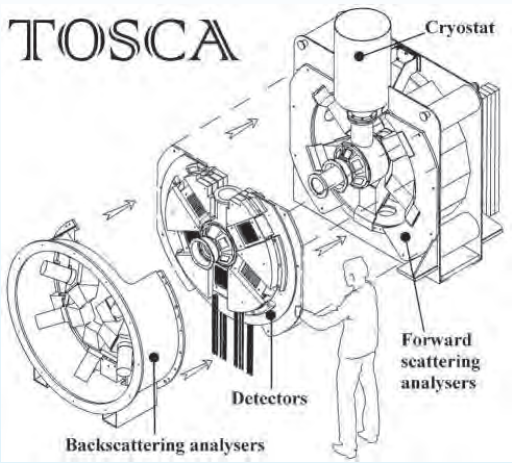
Lattice Dynamics of Crystals

ab initio Lattice Dynamics

Examples

Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

The TOSCA INS Spectrometer



- ISIS pulsed neutron source at Rutherford Appleton Laboratory, UK
- TOSCA - high-resolution powder spectrometer at ISIS
- Little or no anharmonic overtone contamination of spectra
- No selection rule absences
- Signal is *weighted DOS integral* over BZ

- Motivations for *ab initio* lattice dynamics I
- Motivations for *ab initio* lattice dynamics II

Lattice Dynamics of Crystals

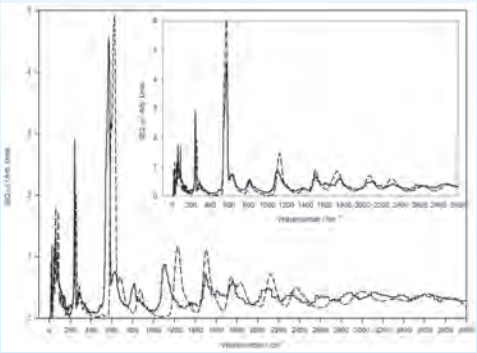
ab initio Lattice Dynamics

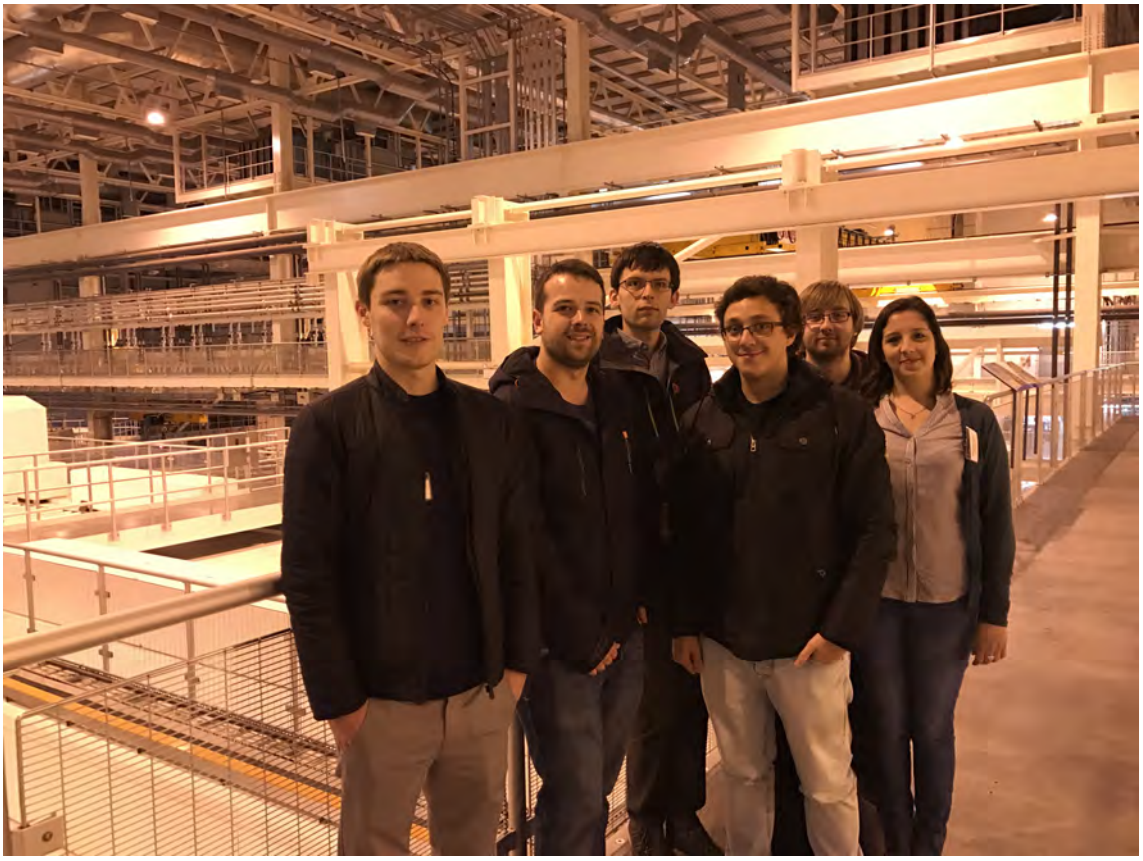
Examples

Modelling of spectra
- Interaction of photons with phonons
- Powder infra-red
- Raman Spectroscopy
- Raman scattering of t-ZrO2
- Inelastic Neutron Scattering
- The TOSCA INS Spectrometer
- INS of Ammonium Fluoride

INS of Ammonium Fluoride

- NH_4F is one of a series of ammonium halides studied in the TOSCA spectrometer. Collab. Mark Adams (ISIS)
- Structurally isomorphic with ice ih
- INS spectrum modelled using A-CLIMAX software (A. J. Ramirez Cuesta, ISIS)
- Predicted INS spectrum in mostly excellent agreement with experiment
- NH_4 libration modes in error by $\approx 5\%$.
- Complete mode assignment achieved.





Simulation-based modeling of neutron scattering data

Linking neutron scattering and MD simulations

Gerald Kneller
Centre de Biophysique Moléculaire, CNRS Orléans
Université d'Orléans
Synchrotron Soleil, St Aubin

Presentation for the MDANSE 2016 school
Abington, UK, 12 November 2016

Content

- Elements of neutron scattering theory
- Linking MD simulations and neutron scattering
- Basic analysis of MD simulations
- Modeling MD time series and correlation functions

Elements of neutron scattering theory

Properties of the neutron

The neutron is a particle of mass $m = 1.6749 \times 10^{-27}$kg with zero charge. A free neutron is not stable and decays after a mean life time of 885.6 s into a proton, an electron, and an electronic anti-neutrino,

$$n \longrightarrow p^+ + e^- + \overline{\nu}_e.$$

A thermal neutron is a non-relativistic article of thermal energy, $E \approx k_B T$. [1] After its generation by nuclear fission or a spallation process, it attains this energy after multiple collisions with the atoms of a moderator (H_20, D_20,...) at temperature T.

Wave length of thermal neutrons

Using the De Broglie relations[2]

$$\vec{p} = \hbar \vec{k}, \qquad |\vec{k}| = \frac{2\pi}{\lambda},$$

one finds for $E = \vec{p}^{\,2}/2m = k_B T$ et $T = 300\,\mathrm{K}$ a wavelength of

$$\lambda = 1.778\,\text{Å}.$$

The wave length is this compatible with typical interatomic distances between the atoms in condensed matter systems. Since $E \approx k_B T$ is comparable with their energy, thermal neutron scattering is a unique tool for studying the structure and the dynamics of condensed matter.

[2] $\hbar = h/(2\pi) = 1.05457 \times 10^{34}$ Js is the reduced Planck constant.

Interaction of neutrons with matter

Neutrons interact with matter primarily through a short-ranged (fm) neutron-nucleus interaction, which is described through Fermi's pseudo-potential,

$$T = \frac{2\pi\hbar^2}{m} b\,\delta(\vec{r} - \vec{R}).$$

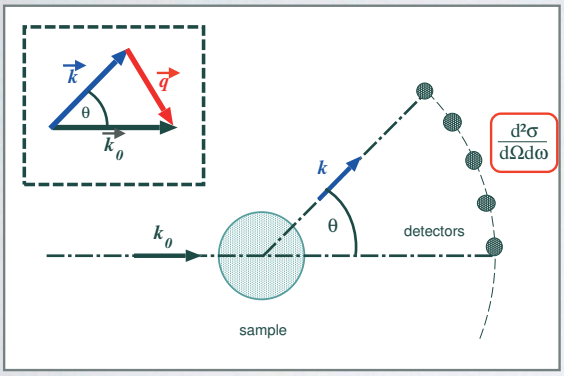
Here $\vec{r}$ and $\vec{R}$ is, respectively, the position operator of the neutron and the nucleaus of the scattering atom and the (generally complex) scattering length b takes values in the fm range. It depends on the relative orientation of the neutron and the nuclear spin. The symbol m denotes the neutron mass. The scattering cross section of a fixed atom is

$$\sigma = 4\pi|b|^2.$$

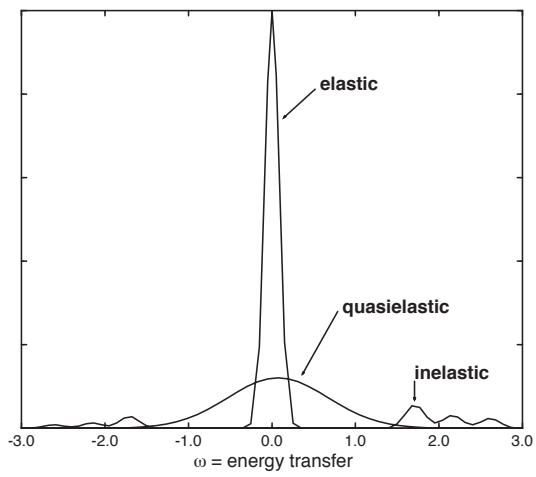
Scattering experiment

Exploring the structural dynamics of condensed matter on the atomic scale (0.1-10 nm, sub ps - 10 ns)

Setup for a neutron scattering experiment



Energy spectrum



Momentum transfer:
$\mathbf{q} = \mathbf{k}_0 - \mathbf{k} = (\mathbf{p}_0 - \mathbf{p})/\hbar$

Energy transfer:
$\omega = (E_0 - E)/\hbar$

Differential scattering cross section

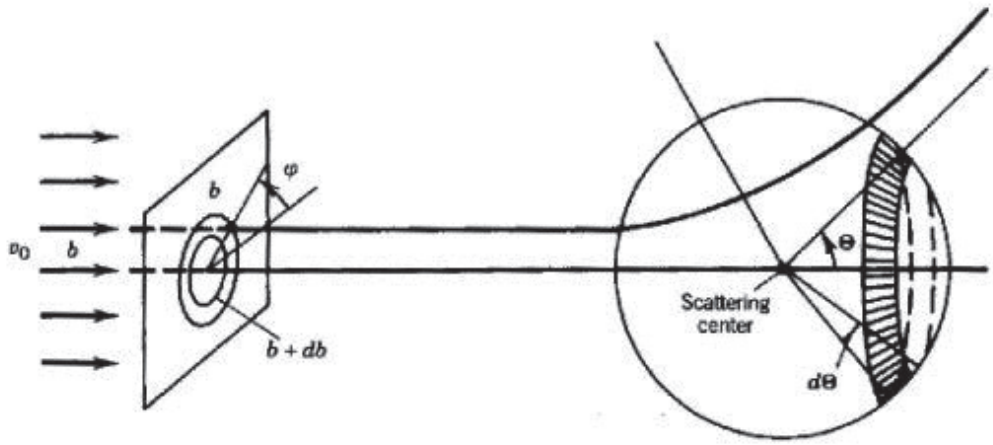


Figure 1-2-5. Relation between the impact parameter b and the CM scattering angle Θ for the one-particle equivalent problem of a purely repulsive Coulomb potential $V_{coul}(r)$. The outer circle in the plane at the left corresponds to the values b and Θ, the inner circle to the values $b - |db|$ and $\Theta + |d\Theta|$. [Adapted from Goldstein (1980).]

The (normalized) differential scattering cross section for N scattering atoms and an unpolarized neutron beam/sample is

$$\boxed{\frac{d^2\sigma}{d\Omega d\omega} = \frac{|\vec{k}|}{|\vec{k}_0|}\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} dt\, e^{-i\omega t}\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\overline{b_\alpha^* b_\beta}\left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)}e^{i\vec{q}\cdot\vec{R}_\beta(t)}\right\rangle}$$

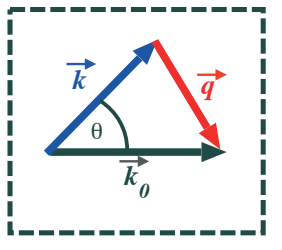
where

$$\boxed{\langle A(0)B(t)\rangle = \frac{1}{Z}\mathrm{tr}\left\{Ae^{iHt/\hbar}Be^{-iHt/\hbar}\right\}}$$

denotes a quantum time correlation function. Here H is the Hamilton operator of the scattering system and $Z = \mathrm{tr}\{e^{-\beta H}\}$ is the partition function, with $\beta = (k_B T)^{-1}$. The overline denotes an average over relative neutron-nucleus spin orientations. The incident and scattered neutron have the momenta $\hbar\vec{k}_0$ and $\hbar\vec{k}$, respectively.

Dynamic structure factor

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{|\vec{k}|}{|\vec{k}_0|}\mathcal{S}(\vec{q},\omega)$$



$$\mathcal{S}(\vec{q},\omega) = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} dt\, e^{-i\omega t}\mathcal{I}(\vec{q},t)$$

Dynamic structure factor

$$\mathcal{I}(\vec{q},t) = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\overline{b_\alpha^* b_\beta}\left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)}e^{i\vec{q}\cdot\vec{R}_\beta(t)}\right\rangle$$

Intermediate scattering function

Coherent and incoherent scattering

The intermediate scattering function is split into a coherent part, reflecting collective motions, and and incoherent part, reflecting single particle motions,

$$\mathcal{I}(\vec{q},t) = \mathcal{I}_{\text{coh}}(\vec{q},t) + \mathcal{I}_{\text{inc}}(\vec{q},t)$$

Defining $b_{\alpha\,\text{coh}} = \overline{b_\alpha}$ and $b_{\alpha\,\text{inc}} = \sqrt{\overline{|b_\alpha|^2} - \left|\overline{b_\alpha}\right|^2}$, one has

$$\mathcal{I}_{\text{coh}}(\vec{q},t) = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} b_{\alpha\,\text{coh}}^* b_{\beta\,\text{coh}}\left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)}e^{i\vec{q}\cdot\vec{R}_\beta(t)}\right\rangle$$

$$\mathcal{I}_{\text{inc}}(\vec{q},t) = \frac{1}{N}\sum_{\alpha=1}^{N}|b_{\alpha\,\text{inc}}|^2\left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)}e^{i\vec{q}\cdot\vec{R}_\alpha(t)}\right\rangle$$

Scattering cross sections for neutrons and photons

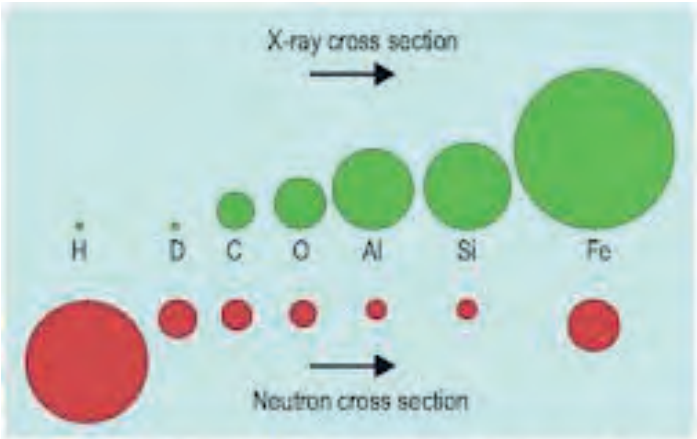


Fig. 2. Neutron and x-ray scattering cross-sections compared. Note that neutrons penetrate through Al much better than x rays do, yet are strongly scattered by hydrogen.

$$\sigma_{tot} = \int\int d\Omega d\omega \frac{d^2\sigma}{d\Omega d\omega}$$

Source: http://www.ncnr.nist.gov/AnnualReport/FY2003_html/RH2/

Hydrogenous samples

Element	H	D	C	O	N	S
b_{coh}	−3.741	6.674	6.648	5.805	9.300	2.847
b_{inc}	25.217	4.022	0.285	0.000	2.241	0.188

Due to the dominant incoherent scattering cross section of hydrogen, soft matter samples (polymers, biomolecules), which contain typically 50 % hydrogen atoms, scatter essentially incoherently. Therefore

$$\boxed{\mathcal{I}(\vec{q},t) \approx \frac{|b_{H\,\mathrm{inc}}|^2}{N_H} \sum_{\alpha\in\mathbb{H}} \left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)} e^{i\vec{q}\cdot\vec{R}_\alpha(t)} \right\rangle .}$$

Classical approximation of the scattering law

In the (usual) classical approximation, quantum time correlation functions are replaced by their classical counterparts, such that

$$\boxed{\mathcal{I}(\vec{q},t) \approx \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\overline{b_\alpha^* b_\beta}\left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)}e^{i\vec{q}\cdot\vec{R}_\beta(t)}\right\rangle_{cl}}$$

Here

$$\langle A(0)B(t)\rangle_{cl} \equiv \frac{1}{Z_{cl}}\int d^{6N}\Gamma e^{-\beta H(\Gamma)}A(\Gamma)e^{t\mathcal{L}}B(\Gamma)$$

is an ensemble average over the $6N$ phase space coordinates Γ, with $Z_{cl} = \int d^{6N}\Gamma e^{-\beta H(\Gamma)}$, and $\mathcal{L} = \sum_{i=1}^{3N}\left\{\frac{\partial H}{\partial p_i}\frac{\partial}{\partial x_i} - \frac{\partial H}{\partial x_i}\frac{\partial}{\partial p_i}\right\}$ is the Liouville operator of the scattering system.

Van Hove functions and QENS

Within the classical approximation, the spatial Fourier transforms of the atomic contributions to the intermediate scattering function,[3]

$$\begin{aligned} G_{\alpha\beta}(\vec{r},t) &\equiv \frac{1}{(2\pi)^3}\int d^3q\, e^{i\vec{q}\cdot\vec{r}}\left\langle e^{-i\vec{q}\cdot\left(\vec{R}_\alpha(0)-\vec{R}_\beta(t)\right)}\right\rangle_{cl}, \\ &= \langle\delta(\vec{r}-[\vec{R}_\alpha(0)-\vec{R}_\beta(t)])\rangle_{cl} \end{aligned}$$

have a simple interpretation: They give the probability to find atom β at time t at position $\vec{r}$, given atom α was at time 0 at position $\vec{0}$. With the above definition

$$\mathcal{I}(\vec{q},t) = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\overline{b_\alpha^* b_\beta}\int d^3r\, e^{-i\vec{q}\cdot\vec{r}}G_{\alpha\beta}(\vec{r},t).$$

[3] L. Van Hove, Physical Review 95, 249 (1954).

[4] L. Van Hove, Physica 24, 404 (1958).

Van Hove function for a freely diffusing atom

In case of free diffusion, the Van Hove self-correlation function fulfills the diffusion equation

$$\partial_t G_s(\mathbf{r}, t) = D \left\{ \partial_x^2 + \partial_y^2 + \partial_z^2 \right\} G_s(\mathbf{r}, t),$$

where D is the diffusion constant. The solution is a Gaussian

$$\boxed{G(\mathbf{r}, t) = \frac{e^{-\frac{|r|^2}{4\mathrm{D}|t|}}}{2\sqrt{\pi \mathrm{D}\,|t|}^3}}$$

The mean square displacement grows linearly with time

$$\boxed{W(t) = \int d^3 r\, |r|^2 G(\mathbf{r}, t) = 6Dt.}$$

Historical articles about diffusion

5. *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen;*
von A. Einstein.

In dieser Arbeit soll gezeigt werden, daß nach der molekularkinetischen Theorie der Wärme in Flüssigkeiten suspendierte Körper von mikroskopisch sichtbarer Größe infolge der Molekularbewegung der Wärme Bewegungen von solcher Größe ausführen müssen, daß diese Bewegungen leicht mit dem Mikroskop nachgewiesen werden können. Es ist möglich, daß die hier zu behandelnden Bewegungen mit der sogenannten „Brownschen Molekularbewegung" identisch sind; die mir erreichbaren Angaben über letztere sind jedoch so ungenau, daß ich mir hierüber kein Urteil bilden konnte.

A. Einstein, *Ann. Phys.*, vol. 322, no. 8, 1905.

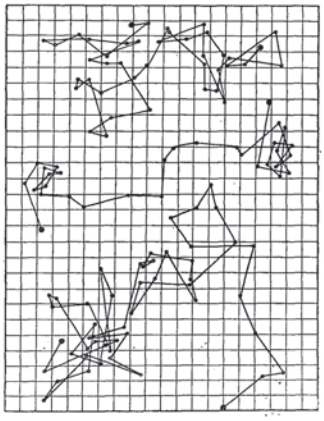
f(x,t) is a concentration

$$f(x, t+\tau)\, dx = dx \cdot \int_{\Delta=-\infty}^{\Delta=+\infty} f(x+\Delta)\, \varphi(\Delta)\, d\Delta \quad \Rightarrow \quad \boxed{\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2}.}$$

$$\boxed{f(x, t) = \frac{n}{\sqrt{4\pi D}} \frac{e^{-\frac{x^2}{4Dt}}}{\sqrt{t}}.} \quad \Rightarrow \quad \boxed{\lambda_x = \sqrt{\overline{x^2}} = \sqrt{2Dt}.}$$

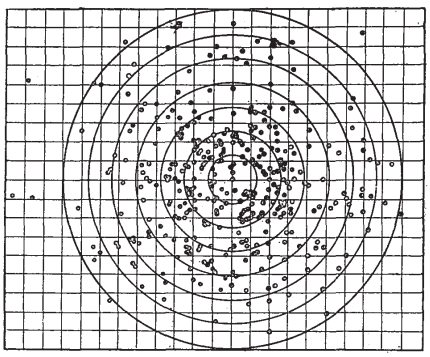
MOUVEMENT BROWNIEN ET RÉALITÉ MOLÉCULAIRE;

PAR M. JEAN PERRIN.



$$\xi^2 = \tau \frac{RT}{N} \frac{1}{3\pi a \zeta}$$

Fig. 7.



Ici encore le contrôle de la loi de répartition peut être quantitatif. Si, en effet, on admet la loi de probabilité donnée pour une composante x, il est facile de voir que la probabilité pour qu'un déplacement horizontal ait une longueur comprise entre r et $r + dr$ est donnée par l'expression

$$\frac{1}{2\pi\xi^2} e^{-\frac{r^2}{2\xi^2}} 2\pi r\, dr,$$

Annales de Chimie et de Physique, vol. 18, p. 5 (1909)

Dynamic structure factor for QENS

The intermediate scattering function for a freely diffusing particle is

$$\mathcal{F}(\mathbf{q}, t) = |b_H|^2 \int d^3r\, e^{i\mathbf{q}\cdot\mathbf{r}} G(\mathbf{r}, t) = |b_H|^2 e^{-\mathrm{D}|q|^2|t|}$$

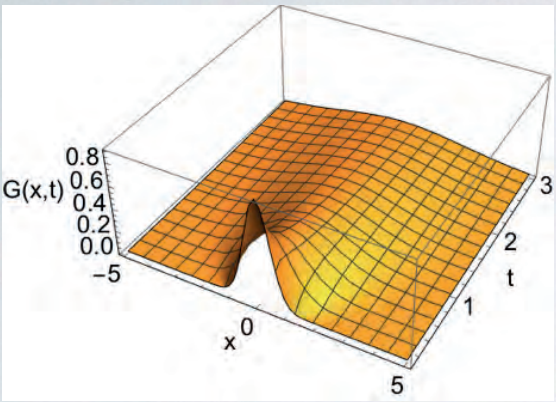
and the corresponding dynamic structure factor

$$\mathcal{S}(\mathbf{q}, \omega) = \frac{|b_H|^2}{2\pi} \int_{-\infty}^{+\infty} dt \int d^3r\, e^{(i\mathbf{q}\cdot\mathbf{r} - i\omega t)} G(\mathbf{r}, t)$$

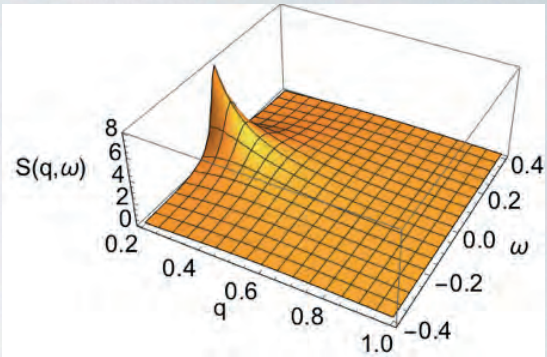
$$= |b_H|^2 \boxed{\frac{\mathrm{D}|q|^2}{\pi\left(\mathrm{D}^2|q|^4 + \omega^2\right)}}$$

Lorentzian quasielastic scattering profile

Van Hove function



Dynamic structure factor



In oder to describe QENS from more complex systems, various diffusion models have been developed (see e.g. [1]).

M. Bée. Quasielastic Neutron Scattering, Adam Hilger, Bristol, 1988.

Limits of the classical approximation in neutron scattering

There are two conditions for the validity of the classical approximation:

1. The intrinsic quantum properties of the scattering system must be negligible.
2. The recoil effect on the scattering atom must be negligible,

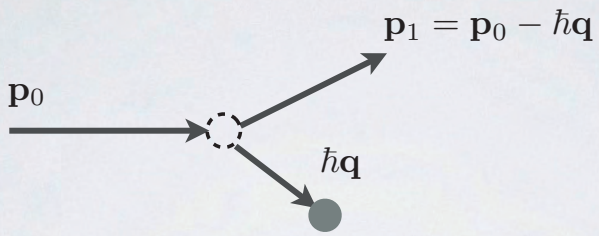
$$\boxed{\frac{\hbar^2|\vec{q}|^2}{2M} \ll k_B T,}$$

where M is the mass of the scattering atom.

1 A. Rahman, Physical Review 130, 1334 (1963).
2 G.R. Kneller, Mol Phys 83, 63 (1994).

Recoil effects

In simulating neutron scattering spectra on the basis of classical molecular dynamics simulations one implicitly assumes the absence of strong recoil effects



Condition for the validity of S(q,w) calculated from MD

$$\frac{\hbar^2 q^2}{2m} \ll k_B T$$

G. Kneller, *Mol. Phys.*, vol. 83, no. 1, pp. 63–87, 1994.

Two very different representations of the same correlation function…

The impact of the neutron on the scattering system can be highlighted through the identity[4]

$$\boxed{\left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)} e^{i\vec{q}\cdot\vec{R}_\alpha(t)} \right\rangle = \left\langle e^{iH_\alpha(\vec{q})t/\hbar} e^{-iHt/\hbar} \right\rangle}$$

where $H_\alpha(\vec{q})$ is the "kicked Hamiltonian"

$$H_\alpha(\vec{q}) = \sum_{\gamma=1}^{N} \frac{(\vec{p}_\gamma + \delta_{\gamma\alpha}\hbar\vec{q})^2}{2M_\gamma} + V(\vec{R}_1, \ldots, \vec{R}_N).$$

One uses here that $e^{-i\vec{q}\cdot\vec{R}_\alpha} e^{iHt/\hbar} e^{i\vec{q}\cdot\vec{R}_\alpha} = e^{iH_\alpha(\vec{q})t/\hbar}$. The usual classical approximation neglects not only quantum effects in the scattering system, but also its perturbation through the neutron.[5]

[4] G. Wick, Physical Review 94, 1228 (1954)
[5] G.R. Kneller, Mol Phys 83, 63 (1994)

Linking MD simulations and neutron scattering

- MD simulations and neutron scattering cover comparable length and time scales.
- Due to the point-like neutron-nucleus interaction, neutron scattering probes the space-time correlations of the same objects which are simulated in MD simulations - the atomic nuclei.
- Compute time correlation functions from MD trajectories, using the classical approximation.

$$\mathcal{I}(\vec{q},t) = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\overline{b_\alpha^* b_\beta}\left\langle e^{-i\vec{q}\cdot\vec{R}_\alpha(0)}e^{i\vec{q}\cdot\vec{R}_\beta(t)}\right\rangle$$

from MD

Time correlation functions from MD simulations

Time correlation functions, such as $\mathcal{I}(\vec{q},t)$, can be computed from time series, assuming ergodicity ($t \geq 0$),

$$\langle A(0)B(t)\rangle_{cl} = \lim_{T\to\infty}\frac{1}{T-t}\int_0^{T-t} d\tau\, A(\tau+t)B(\tau).$$

For discrete time series (MD trajectories) one uses the approximation ($m = 0, 1, 2, \ldots$)

$$\langle A(0)B(m)\rangle_{cl} \approx \frac{1}{N_t - m}\sum_{k=0}^{N_t-m-1} A(k+m)B(k),$$

where N_t is the number of available time frames in the series.

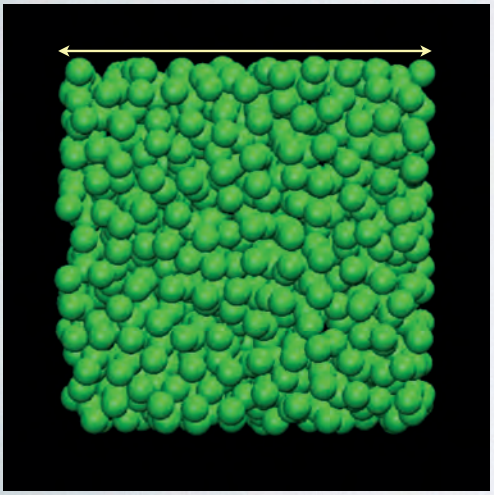
Principle of MD simulations

PHYSICAL REVIEW VOLUME 136, NUMBER 2A 19 OCTOBER 1964

Correlations in the Motion of Atoms in Liquid Argon*

A. RAHMAN
Argonne National Laboratory, Argonne, Illinois
(Received 6 May 1964)

~ 3.6 nm



- Solve Newton's equation of motion

$$M_i \ddot{\mathbf{r}}_i = -\frac{\partial U}{\partial \mathbf{r}_i} \qquad U = \sum_{ij} 4\epsilon \left(\left[\frac{\sigma}{r_{ij}}\right]^{12} - \left[\frac{\sigma}{r_{ij}}\right]^{6} \right)$$

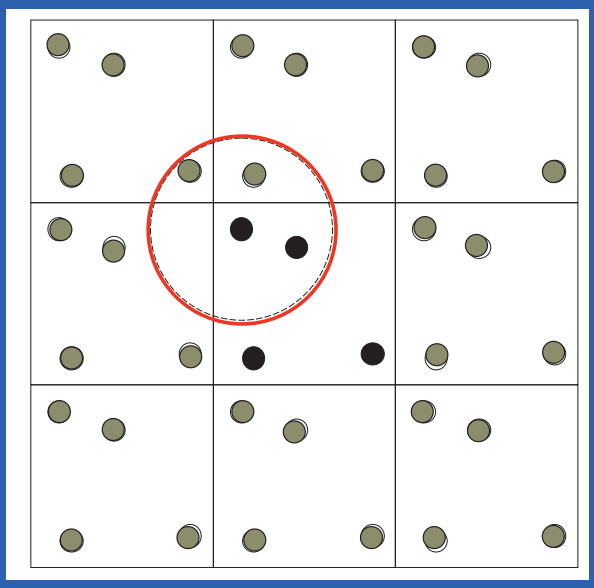
- Discretization and iterative solution itérative yields trajectories = time series (< 100 ns)

$$\mathbf{r}_i(n+1) \leftarrow 2\mathbf{r}_i(n) - \mathbf{r}_i(n-1) + \frac{\Delta t^2}{M_i}\mathbf{F}_i(n)$$

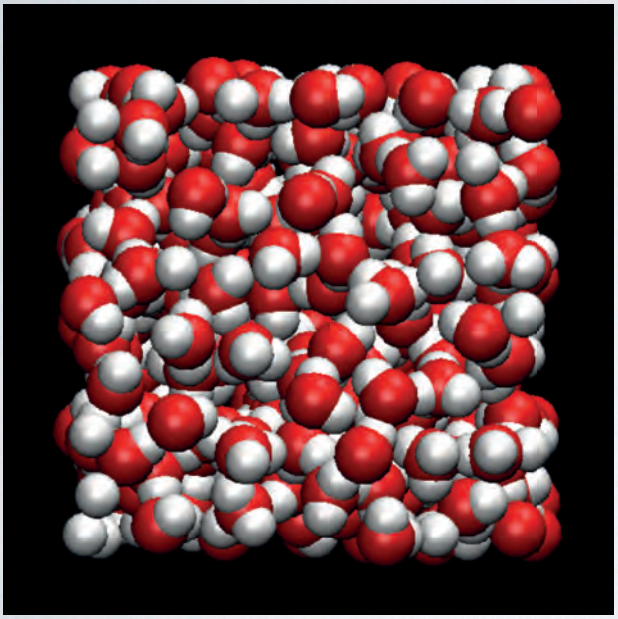
$$\mathbf{v}_i(n) \leftarrow \frac{\mathbf{r}_i(n+1) - \mathbf{r}_i(n-1)}{2\Delta t}.$$

Forces: $\mathbf{F}_i = -\frac{\partial U}{\partial \mathbf{r}_i}$

Periodic boundary conditions



Dynamics of water



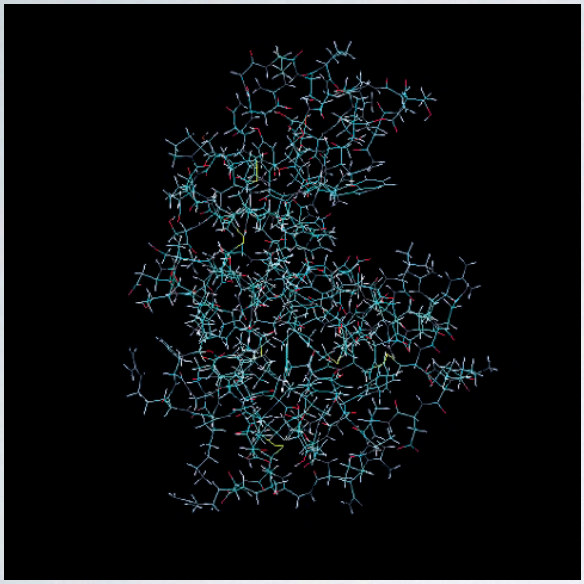
Dynamics of 256 water molecules with in a cibic box with periodic boundary conditions and Ewald summation for the Coulomb forces

SPC/E potential [1]:

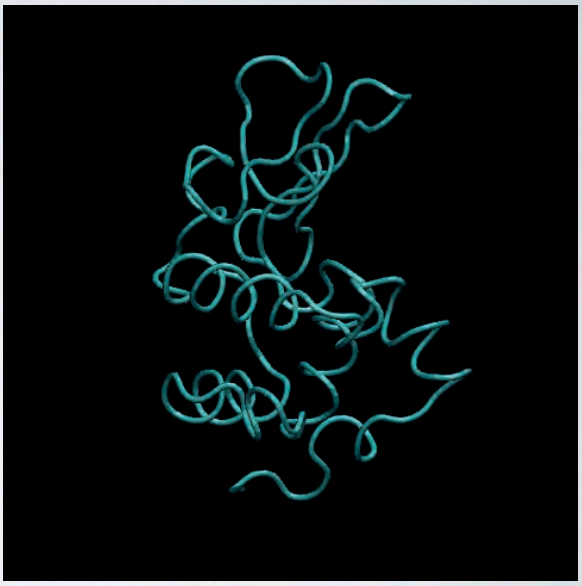
- O-O interactions of Lennard-Jones type
- **Coulomb interactions** for O-O, H-H, O-H

[1] H. Berendsen, J.R. Grigera, and T.P. Straatsma, J Phys Chem-Us 91, 6269 (1987).

Dynamics of lysozyme

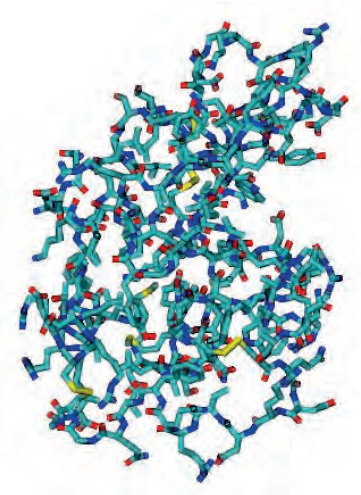


All atoms



backbone

Force field for biomolecular simulations



Lysozyme

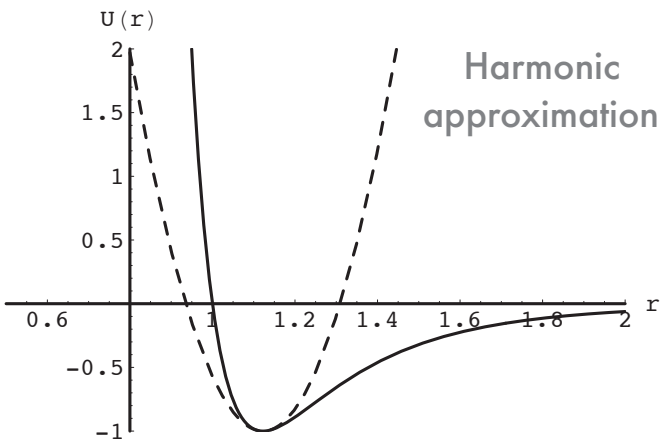
$$U = \sum_{\text{bonds } ij} k_{ij}(r_{ij} - r_{ij}^{(0)})^2 + \sum_{\text{angles } ijk} k_{ijk}(\phi_{ijk} - \phi_{ijk}^{(0)})^2$$

$$+ \sum_{\text{dihedrals } ijkl} k_{ijkl}\cos(n_{ijkl}\theta_{ijkl} - \delta_{ijkl})$$

$$\left.\begin{aligned} &+ \sum_{\text{all pairs } ij} 4\epsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r^{12}} - \frac{\sigma_{ij}^{6}}{r^{6}}\right) \\ &+ \sum_{\text{all pairs } ij} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \end{aligned}\right\} \text{non-bonded.}$$

The force field (Amber)

The limit of classical MD



Harmonic approximation

$$\boxed{\hbar\omega_0 \ll k_B T}$$

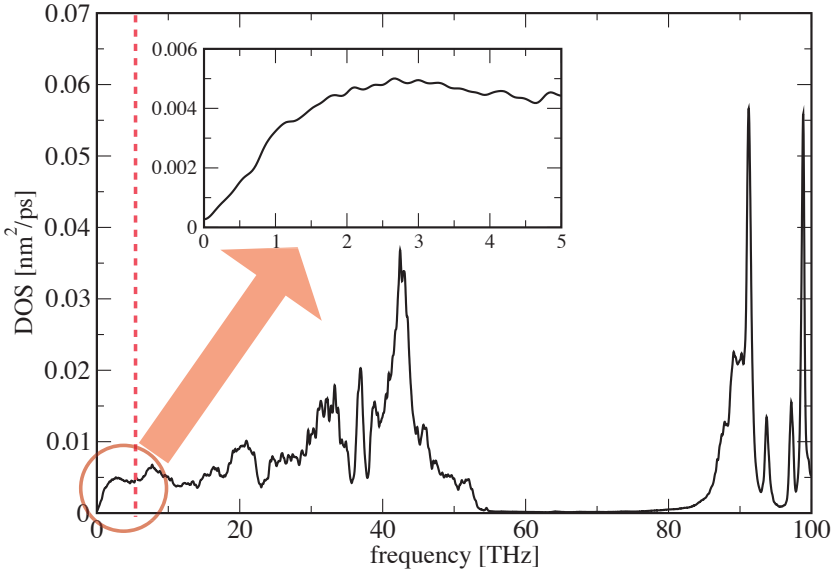
Argon :

$$T = 94.4\,K$$
$$\hbar\omega_0 = 0.4\,k_B T$$

$$U_{LJ}(r) = 4\epsilon\left(\left[\frac{\sigma}{r}\right]^{12} - \left[\frac{\sigma}{r}\right]^{6}\right) \approx -\epsilon + \frac{18\cdot 2^{2/3}\epsilon\left(r - r_0\right)^2}{\sigma^2}$$

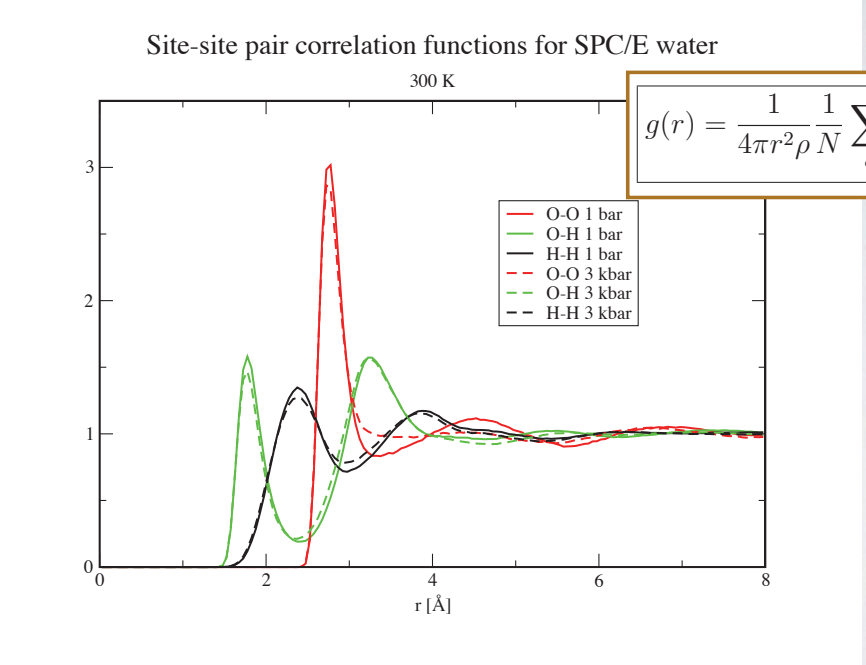
$$\omega_0 = \sqrt{\frac{18\cdot 2^{2/3}\epsilon}{\mu\sigma^2}}$$

For macromolecules like proteins only a small part of the motional frequency spectrum is accessible to classical MD!



Basic analysis of MD simulations

Static site pair correlation functions of water

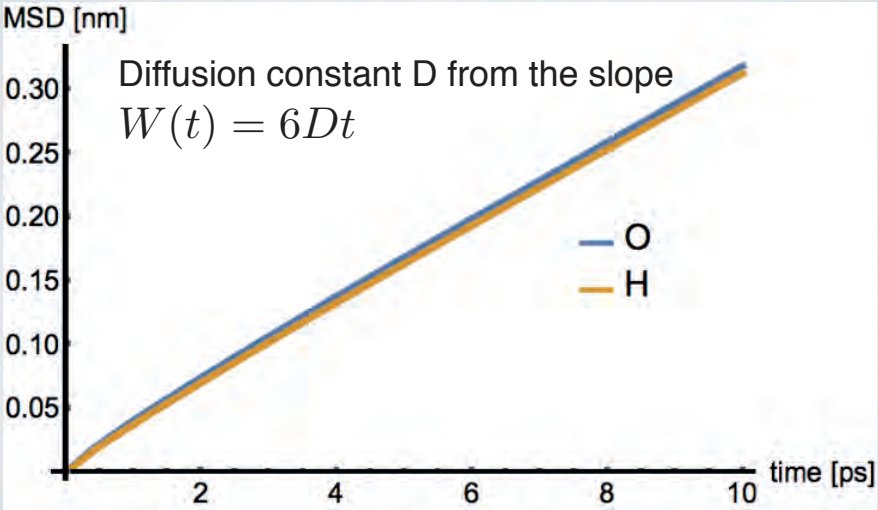


$$g(r) = \frac{1}{4\pi r^2 \rho} \frac{1}{N} \sum_{\alpha} \sum_{\beta \neq \alpha} \langle \delta(r - |R_\alpha - R_\beta|) \rangle$$

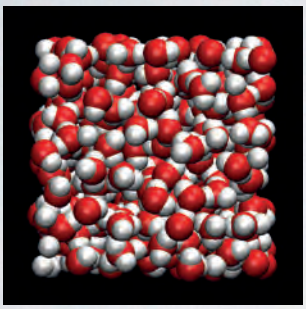
Atomic mean square displacements

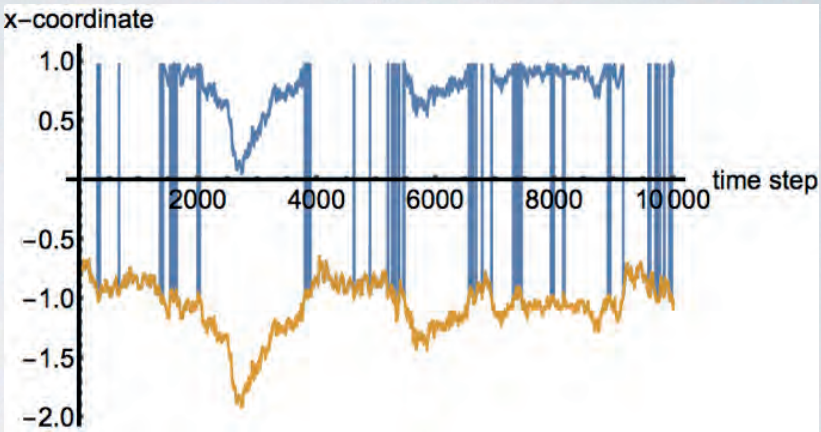
$$W^{(\alpha)}(n) \approx \frac{1}{N_t - |n|} \sum_{k=0}^{N_t-|n|-1} \left| \mathbf{x}^{(\alpha)}(k+n) - \mathbf{x}^{(\alpha)}(k) \right|^2$$

$$W(n) = \frac{1}{N} \sum_{\alpha=1}^{N} W_{vv}^{(\alpha)}(n)$$ average over molecules



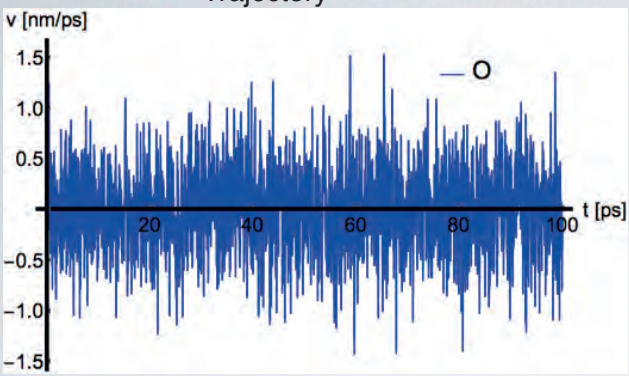
Remove first jumps due to p.b.c. !



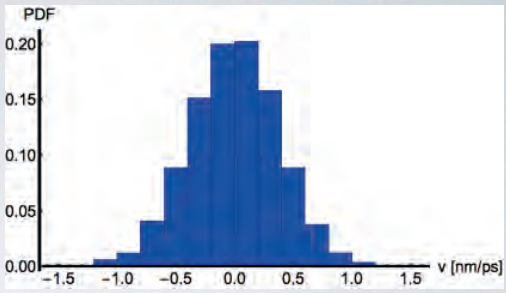


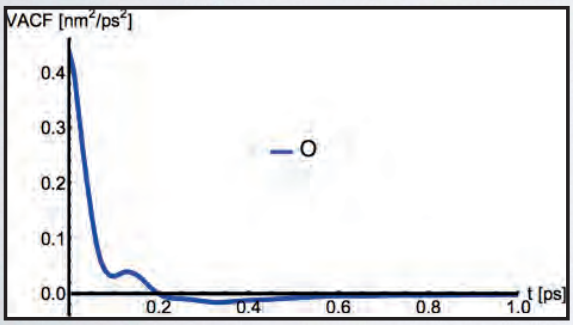
Velocity autocorrelation function (oxygen)

Trajectory



Maxwell distribution





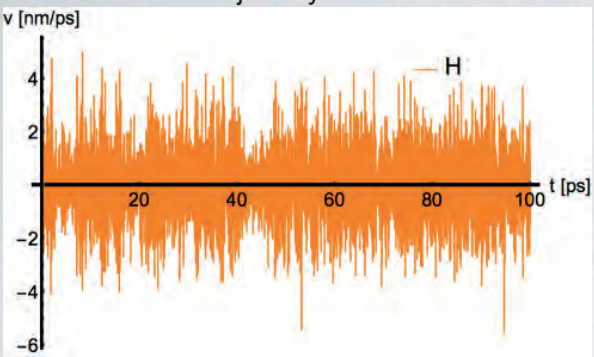
$$c_{vv}^{(\alpha)}(n) \approx \frac{1}{N_t - |n|} \sum_{k=0}^{N_t-|n|-1} \mathbf{v}^{(\alpha)}(k) \cdot \mathbf{v}^{(\alpha)}(k+n)$$

$$c_{vv}(n) = \frac{1}{N} \sum_{\alpha=1}^{N} c_{vv}^{(\alpha)}(n)$$

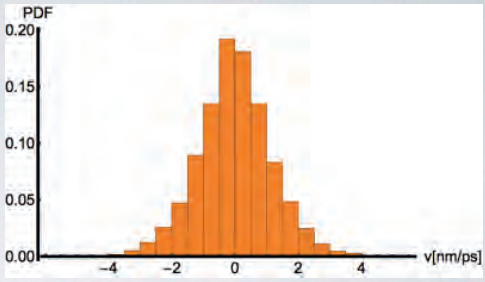
average over molecules

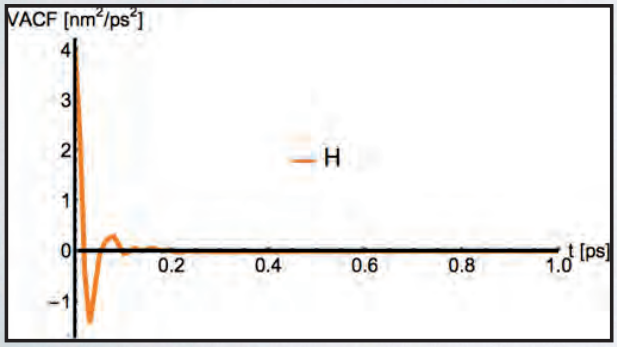
Velocity autocorrelation function (hydrogen)

Trajectory



Maxwell distribution

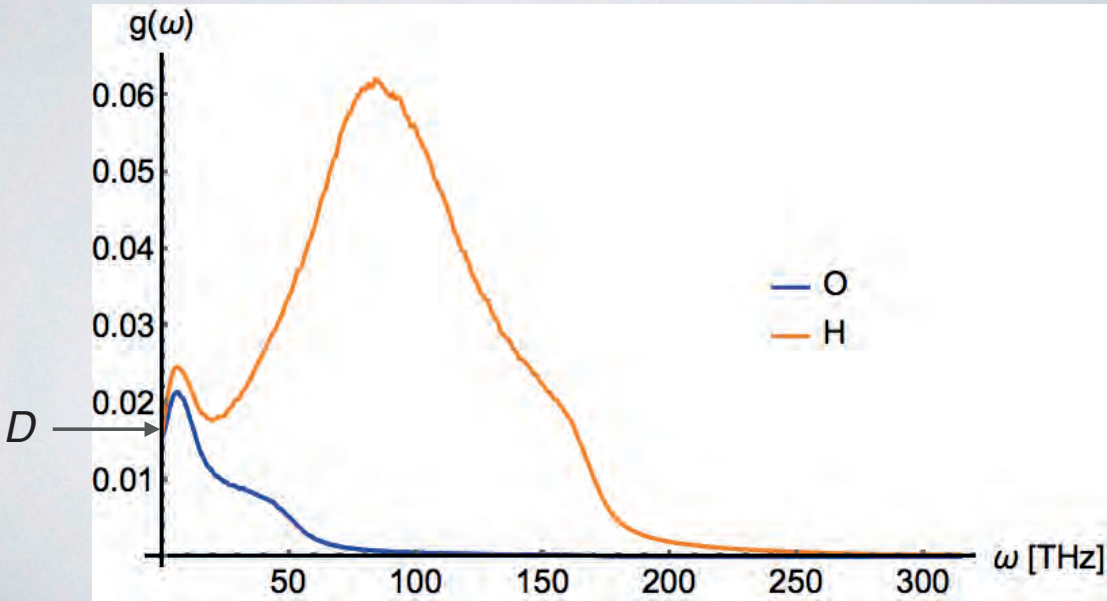




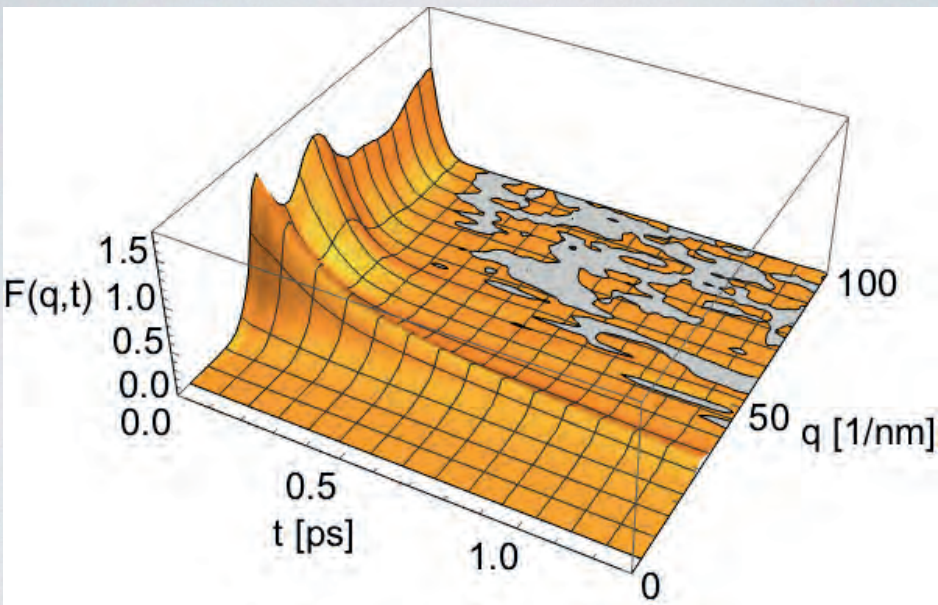
Density of States, VACF, and diffusion coefficient

$$g(\omega) \equiv \int_0^\infty dt\, \cos\omega t\, c_{vv}(t).$$

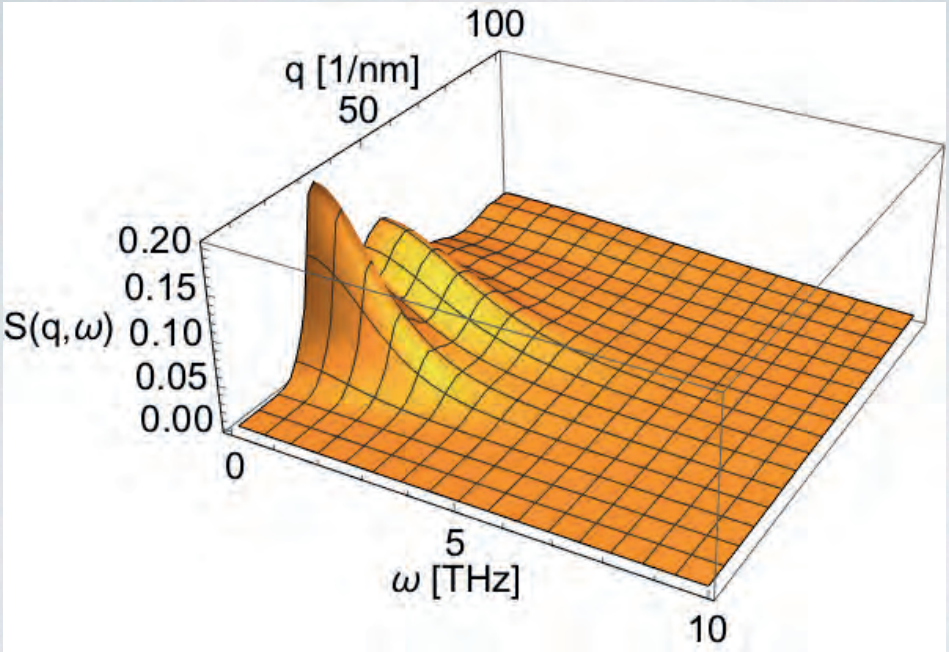
$$D = \int_0^\infty dt\, c_{vv}(t) = g(0),$$



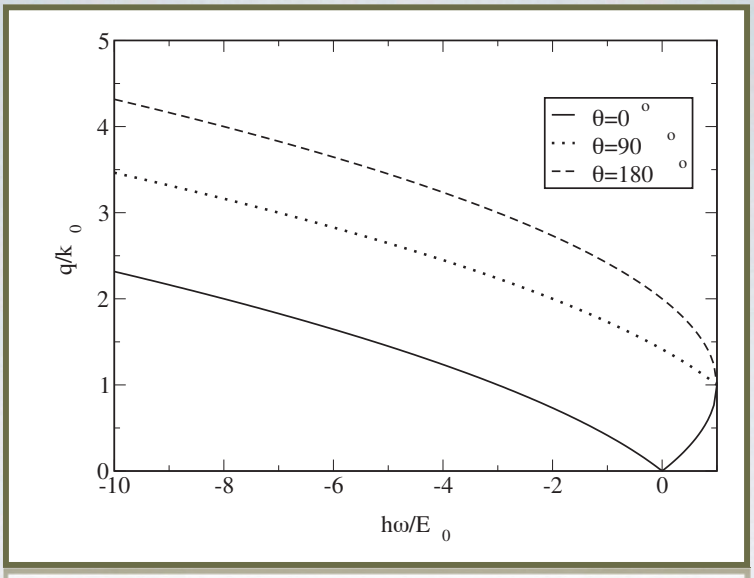
Coherent intermediate scattering function



Coherent dynamic structure factor



Experimentally accessible (q,ω)-range



$$|\vec{q}| = k_0\sqrt{2 - \frac{\hbar\omega}{E_0} - 2\sqrt{1 - \frac{\hbar\omega}{E_0}}\cos\theta}$$

Calculation of time correlation functions - an « efficient detour » via Fourier space

The correlation theorem of the Fourier transform reads

$$\int_{-\infty}^{+\infty} dt\, f(t+\tau)g^*(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega\, \tilde{f}(\omega)\tilde{g}^*(\omega)e^{i\omega t}$$

where

$$\tilde{f}(\omega) = \int_{-\infty}^{+\infty} d\omega\, f(t)e^{-i\omega t},$$
$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega\, \tilde{f}(\omega)e^{i\omega t},$$

is the Fourier transform pair of f.

Fast Correlation algorithm

The discrete version of the correlation theorem,

$$\sum_{k=0}^{N_t-1} f(n+k)g^*(k) = \frac{1}{N_t}\sum_{k=0}^{N_t-1} F(k)G^*(k)e^{2\pi i\frac{nk}{N_t}}$$
$$F(k) \overset{FFT}{=} \sum_{n=0}^{N_t-1} f(n)e^{-2\pi i\frac{nk}{N_t}},$$

can be used to compute correlation functions by a "detour" via a Fast Fourier Transform (FFT), which reduces the complexity from N_t^2 to $N_t \log_2 N_t$, N_t being the number of frames in the discrete signals. To avoid spurious correlations due to periodicity, use zero padding,

$$f(n) \to f_p(n) = \begin{cases} f(n) & \text{if} \quad 0 \le n \le N_t - 1, \\ 0 & \text{if} \quad N_t \le n \le 2N_t - 1. \end{cases}$$

Computing MSDs

Use here that [6]

$$\left\langle \left(x(n+k)-x(k)\right)^2 \right\rangle = 2\underbrace{\langle x(n+k)x(k)\rangle}_{\text{FCA}} + \underbrace{\langle x(n+k)^2\rangle}_{\text{recursively}} + \underbrace{\langle x(k)^2\rangle}_{\text{recursively}}.$$

[6] G.R. Kneller *et al.*, Comput Phys Commun 91, 191 (1995).

Modeling MD time series and correlation functions

Autoregressive time series model

$v(n) \equiv v(n\Delta t), \quad n \in \mathbb{Z}.$ **time series**

$$\boxed{v(n) = \sum_{k=1}^{P} a_k^{(P)} v(n-k) + \epsilon_P(n)}$$

AR model of order P

$$\begin{aligned} \langle \epsilon_P(n) \rangle &= 0, \\ \langle \epsilon_P(n)\epsilon_P(n') \rangle &= \sigma_P^2 \delta_{nn'}. \end{aligned}$$

"white noise"

parameters of the model: $\alpha_1^{(P)}, \ldots, \alpha_P^{(P)}, \sigma_P$

Wiener-Hopf equations for the coefficients

$$\langle \epsilon_P(n) v(n-k) \rangle = 0 \qquad (k = 1, \ldots, P)$$

$$\boxed{\sum_{k=1}^{P} c_{vv}(|j-k|) a_k^{(P)} = c_{vv}(j), \qquad j = 1 \ldots P}$$

yields the coefficients $a_k^{(P)}$

$$\boxed{\sigma_P^2 = c_{vv}(0) - \sum_{k=1}^{P} a_k^{(P)} c_{vv}(k)}$$

Wiener-Khintchin theorem for discrete signals

- Finite sample of a signal $\quad v_M(n) = \begin{cases} v(n) & \text{si} \quad -M \leq n \leq M \\ 0 & \text{sinon} \end{cases}$

- z-Transform

$$f(n) = \frac{1}{2\pi i} \oint_C dz\, z^{n-1} F_{(>)}(z) \longleftrightarrow \begin{aligned} F(z) &= \sum_{n=-\infty}^{+\infty} f(n) z^{-n}. \\ F_>(z) &= \sum_{n=0}^{\infty} f(n) z^{-n}. \end{aligned}$$

$$(f \circ g)(n) = \sum_{j=-\infty}^{+\infty} f(n+j) g^*(j) \longleftrightarrow F(z) G^*(1/z^*)$$

- Correlation function $\quad c_{vv}(n) = \lim_{M \to \infty} \frac{1}{2M+1} \sum_{k=-M}^{M} v(n+k) v^*(k)$

$$\boxed{C_{vv}(z) = \lim_{M \to \infty} \frac{1}{2M+1} V_M(z) V_M^*(1/z^*)}$$

Analytical form of the z-transformed correlation function

AR model

$$V(z) = \frac{\mathcal{E}_P(z)}{1 - \sum_{k=1}^{P} a_k^{(P)} z^{-k}}$$

$$\lim_{M\to\infty} \frac{1}{2M+1} V(z) V^*(1/z^*) = \frac{\lim_{M\to\infty} \frac{1}{2M+1} \mathcal{E}_P(z) \mathcal{E}_P^*(1/z^*)}{\left(1 - \sum_{k=1}^{P} a_k^{(P)} z^{-k}\right)\left(1 - \sum_{l=1}^{P} a_l^{(P)} z^{l}\right)}$$

$$C_{vv}^{(AR)}(z) = \frac{C_{\epsilon\epsilon}(z)}{\left(1 - \sum_{k=1}^{P} a_k^{(P)} z^{-k}\right)\left(1 - \sum_{l=1}^{P} a_l^{(P)} z^{l}\right)}$$

$$C_{vv}^{(AR)}(z) = \frac{\sigma_P^2}{\left(1 - \sum_{k=1}^{P} a_k^{(P)} z^{-k}\right)\left(1 - \sum_{l=1}^{P} a_l^{(P)} z^{l}\right)}$$

"all pole" model

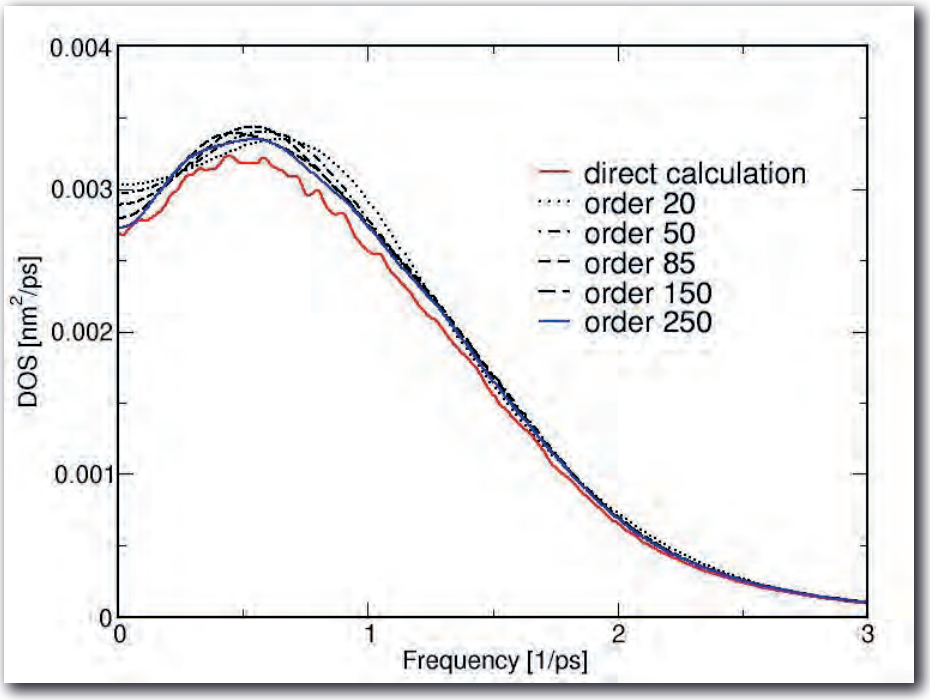
Analytical form of the Fourier-transformed correlation function

$$\tilde{c}_{vv}^{(AR)}(\omega) = \Delta t \sum_{n=-\infty}^{+\infty} c_{vv}^{(AR)}(n) \exp[-in\omega\Delta t] \approx \tilde{c}_{vv}(\omega)$$

$$\tilde{c}_{vv}^{(AR)}(\omega) = \Delta t \, C_{vv}^{(AR)}\left(\exp[i\omega\Delta t]\right)$$

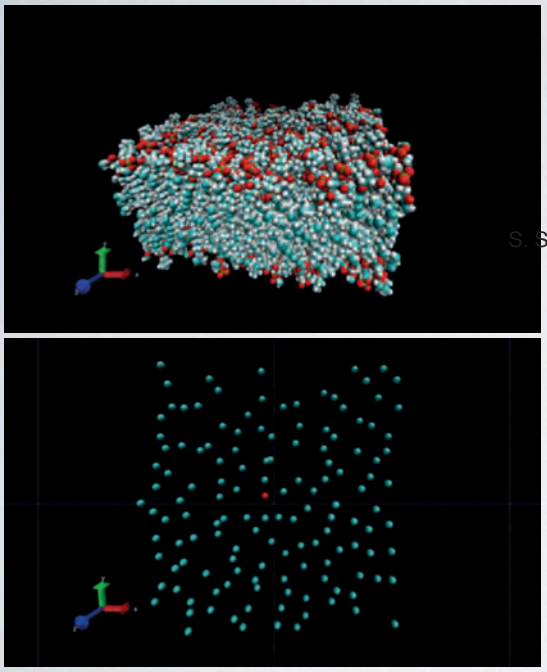
$$g(\omega) \approx \frac{\sigma_P^2 \Delta t}{2\left|1 - \sum_{k=1}^{P} a_k^{(P)} \exp(-i\omega k \Delta t)\right|^2}.$$

Application to a simple liquid (argon)



G.R. Kneller and K. Hinsen. *J. Chem. Phys.*, 115(24):11097–11105, 2001.

Anomalous diffusion of molecules in a lipid bilayer

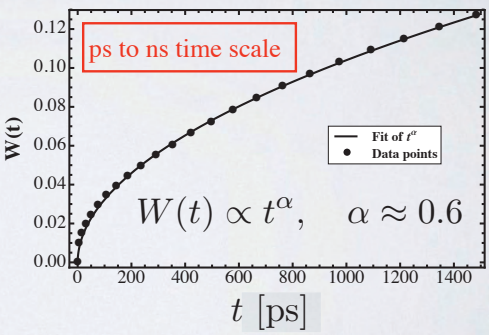


S. Stachura and G.R. Kneller, Mol Sim. 40, 245 (2013).
S. Stachura and G.R. Kneller, J Chem Phys in rint .

- 2x137 POPC molecules (10 nm × 10 nm in the XY-plane)
- 10471 water molecules (fully hydrated)
- OPLS force field
- T=310 K

S. Stachura and G. R. Kneller. Probing anomalous diffusion in frequ

MSD for lateral diffusion



See also

E. Flenner, J. Das, M. Rheinstädter, and I. Kosztin, Phys Rev E 79, 11907 (2009).
G.R. Kneller, K. Baczynski, and M. Pasenkiewicz-Gierula, J Chem Phys 135, 141105 (2011).
J.H. Jeon, H. Monne, M. Javanainen, and R. Metzler, Phys Rev Lett (2012).

Anomalous diffusion is known since long time

ANOMALOUS DIFFUSION IN TRUE SOLUTION.

By Herbert Freundlich and Deodata Krüger.

Received 30th April, 1935.

H. Freundlich and D. Krüger, Trans. Faraday Soc. 31, 906 (1935).

$$\sigma^2(t) := \frac{\int d^n r\, |\mathbf{r}|^2 f(\mathbf{r},t)}{\int d^n r\, f(\mathbf{r},t)}$$

Anomalous Diffusion of Acetone into Cellulose Acetate*

F. A. Long, E. Bagley, and J. Wilkens

Department of Chemistry, Cornell University, Ithaca, New York

(Received May 18, 1953)

F.A. Long, E. Bagley, and J. Wilkens,
The Journal of Chemical Physics 21, 1412 (1953).

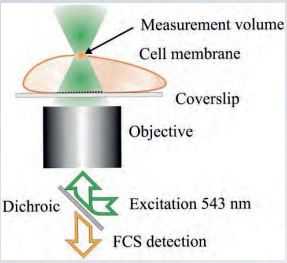
$$\boxed{\sigma^2(t) \propto t^\alpha}$$

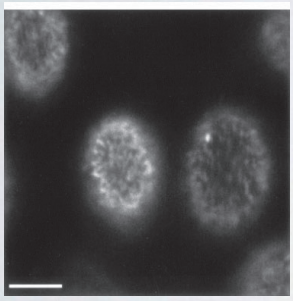
$0 < \alpha < 1$

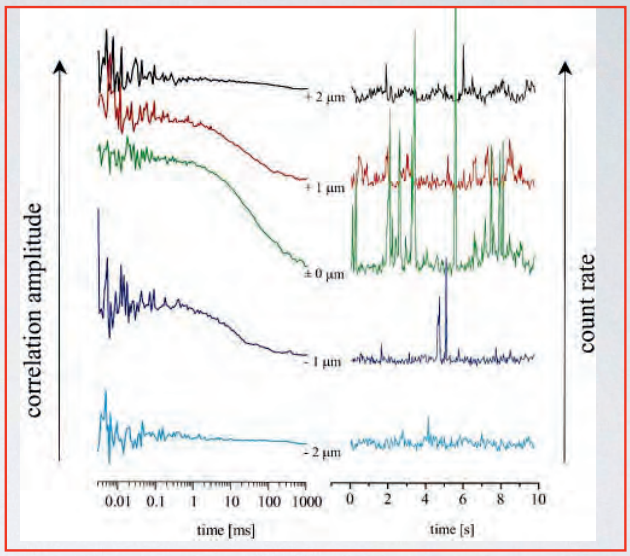
(subdiffusion)

Subdiffusion of lipids observed by FCS

P. Schwille, J. Korlach, and W. Webb, Cytometry 36, 176 (1999).







$$\boxed{\sigma^2(t) \propto t^\alpha}$$

ms to s time scale

Fractional diffusion/Fokker Planck equation

W. Wyss, Journal of Mathematical Physics 27, 2782 (1986).
R. Metzler, E. Barkai, and J. Klafter, Phys Rev Lett 82, 3563 (1999).

$$\frac{\partial}{\partial t} p(\mathbf{r}, t|\mathbf{r}_0, 0) = \partial_t^{1-\alpha} D_\alpha \Delta p(\mathbf{r}, t|\mathbf{r}_0, 0)$$

$$\partial_t^{1-\alpha} g(t) = \frac{d}{dt} \int_0^t d\tau \, \frac{(t-\tau)^{\alpha-1}}{\Gamma(\alpha)} g(\tau)$$

Fractional derivative

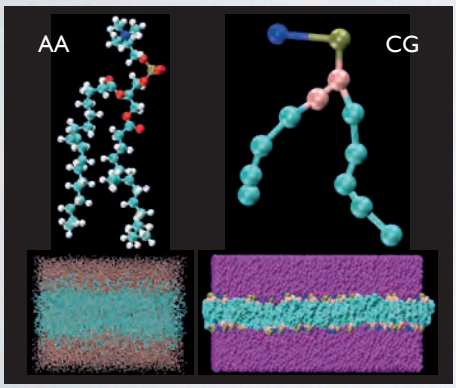
$$W(t) = \frac{2nD_\alpha t^\alpha}{\Gamma(1+\alpha)}.$$

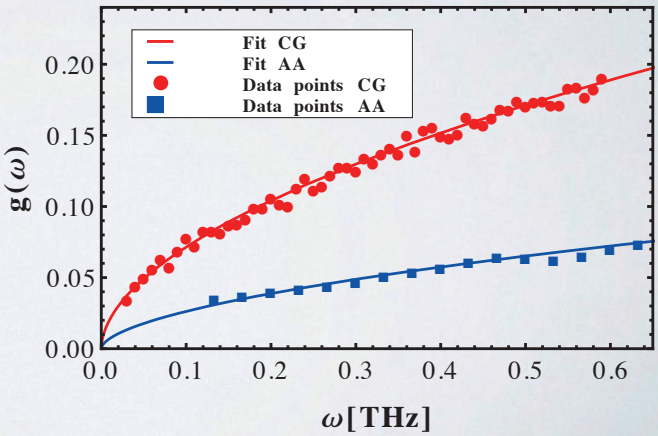
Precise calculation of the low-frequency DOS

S. Stachura and G.R. Kneller, J .Chem. Phys., vol. 143, p. 191103, 2015.

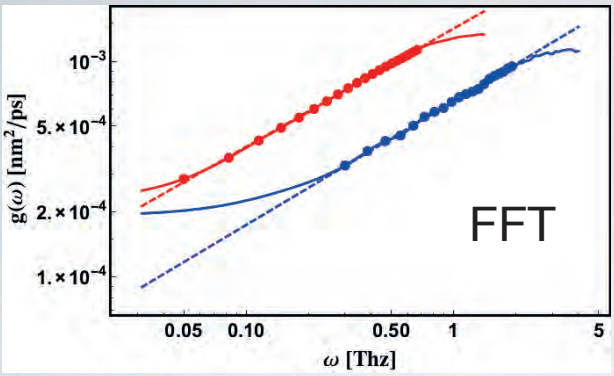
$$g(\omega) = \int_0^\infty dt \, \cos(\omega t) c_{vv}(t) \overset{\omega \ll 1/\tau_v}{\sim} D_\alpha \, \omega^{1-\alpha} \sin\left(\frac{\pi\alpha}{2}\right)$$

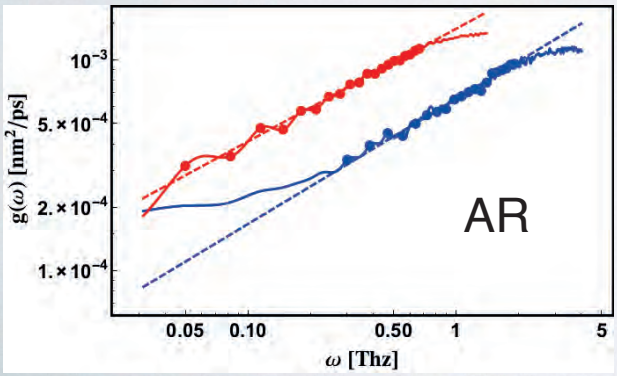
Compare the low freq. DOS for POPC simulations with an all-atom (OPLS) and a coarse-grained (MARTINI) force field:





Low frequency DOS - FFT versus AR estimation

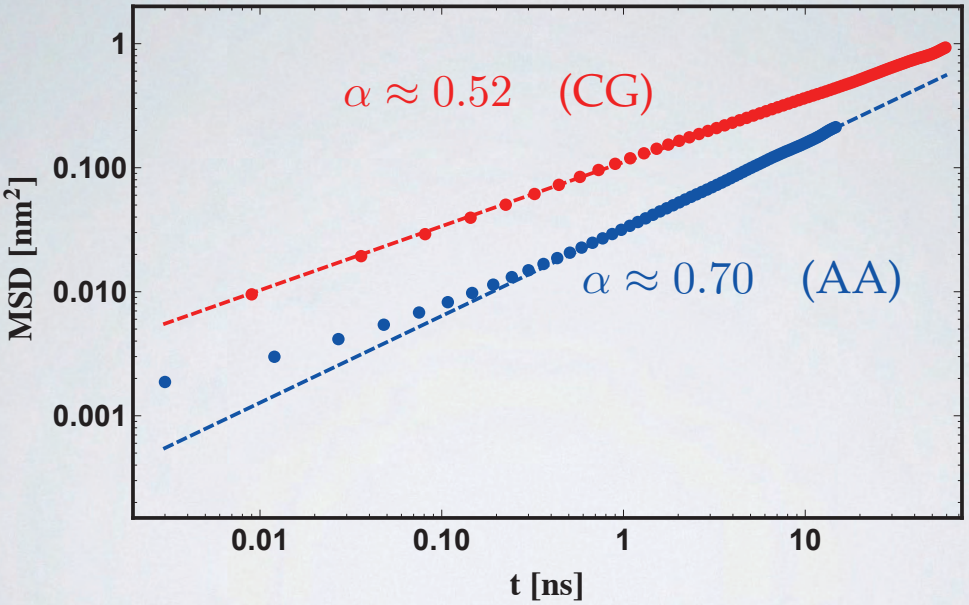




Precision in the lose-frequency regime matters!

Both FFT and AR estimations give very similar results

Mean square displacement (log-log)



α	MSD	WDFT	MEE
AA	0.700 ± 0.003	0.426 ± 0.007	0.406 ± 0.018
CG	0.516 ± 0.002	0.452 ± 0.003	0.466 ± 0.012

D_α	MSD	WDFT	MEE
AA	0.0160 ± 0.0001	0.0225 ± 0.0003	0.0205 ± 0.0007
CG	0.0555 ± 0.0003	0.0466 ± 0.0004	0.0394 ± 0.0012

TABLE I: Fit parameters α and D_α [$\mathrm{nm}^2/\mathrm{ps}^\alpha$], obtained by fits of a) Expression (1) to the MSD, b) Expression (5) to $g(\omega)$ from a windowed discrete Fourier transform, c) Expression (5) to $g(\omega)$ from maximum entropy estimation. Here AA stands for "all atom" and CG to "coarse-grained.

Relation with QENS

Use that $g(\omega) = \pi \lim_{q \to 0} \frac{\omega^2}{q^2} S(q, \omega)$

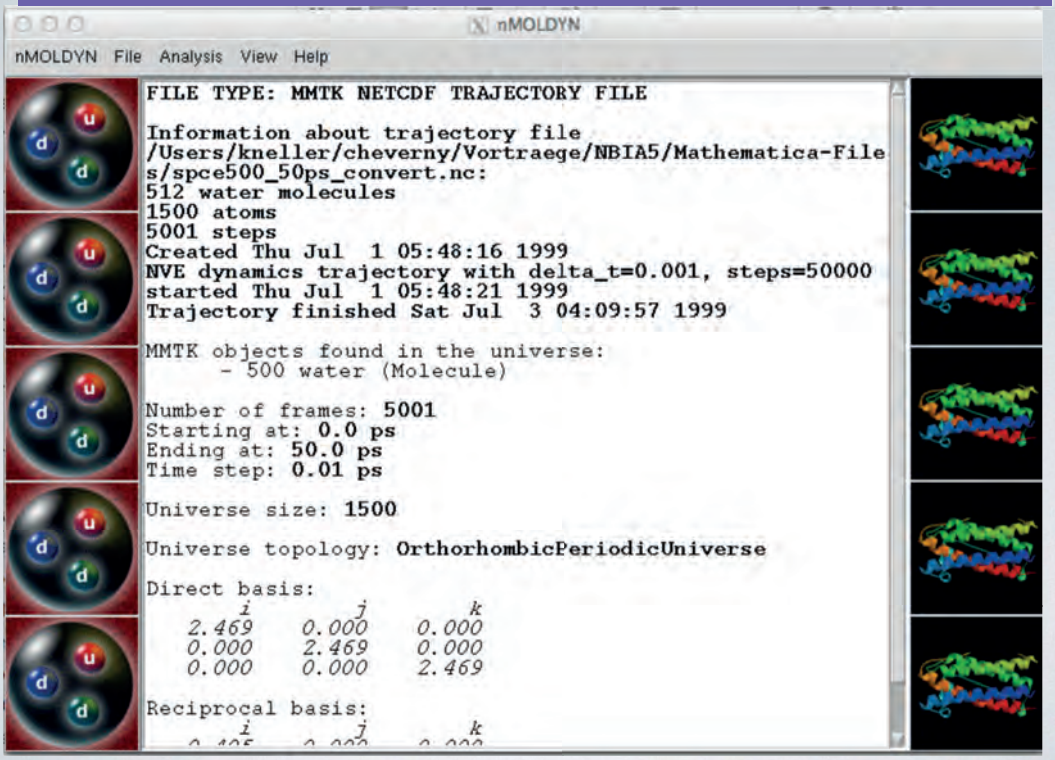
where $g(\omega) \overset{\omega \to 0}{\sim} \omega^{1-\alpha} \sin\left(\frac{\pi\alpha}{2}\right) \Gamma(1+\alpha) D_\alpha$

The asymptotic small frequency regime can be estimated through

$$\omega\tau_v \ll 1, \quad \text{with } \tau_v = \left(\frac{D_\alpha}{\langle|\mathbf{v}^2|\rangle},\right)^{\frac{1}{2-\alpha}}$$

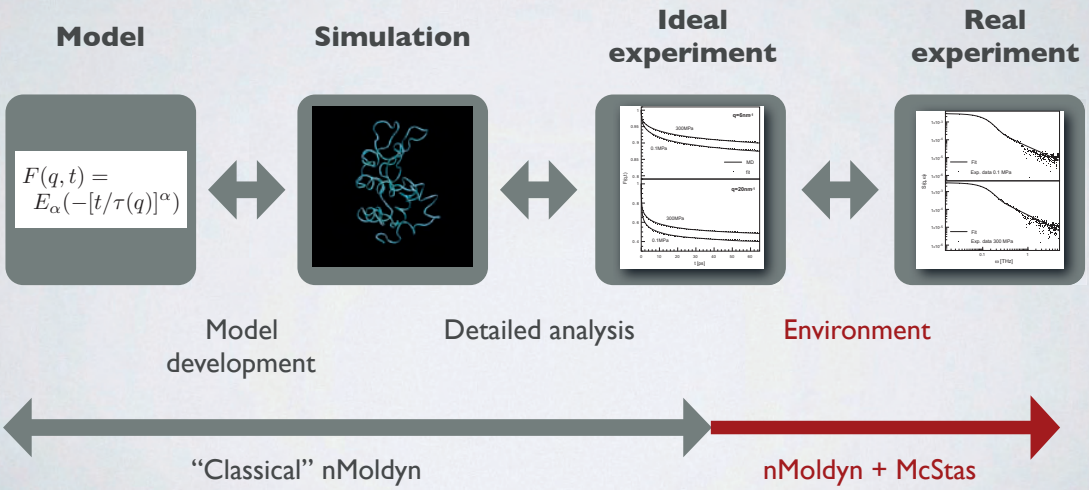
where τ_v is in the picosecond regime.

nMoldyn



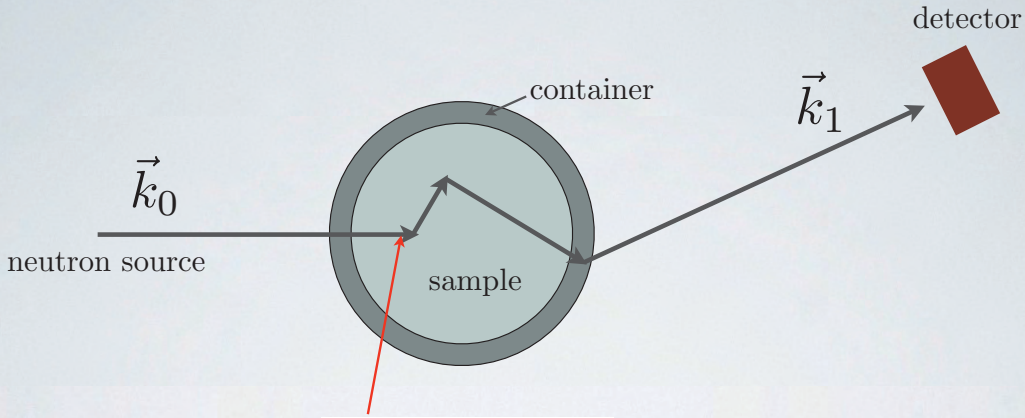
Ideal and «real» *in silico* experiments...

Integrating molecular simulation in "virtual" experiments probing the structure and dynamics of condensed matter



Multiple scattering: example for an undesirable effect

In the MDANSE version of nMoldyn, such effects can be simulated with the McStas instrument simulator



$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{k}{k_0} S(\mathbf{q}, \omega)$$

Differential scatterung cross section for ideal experiment

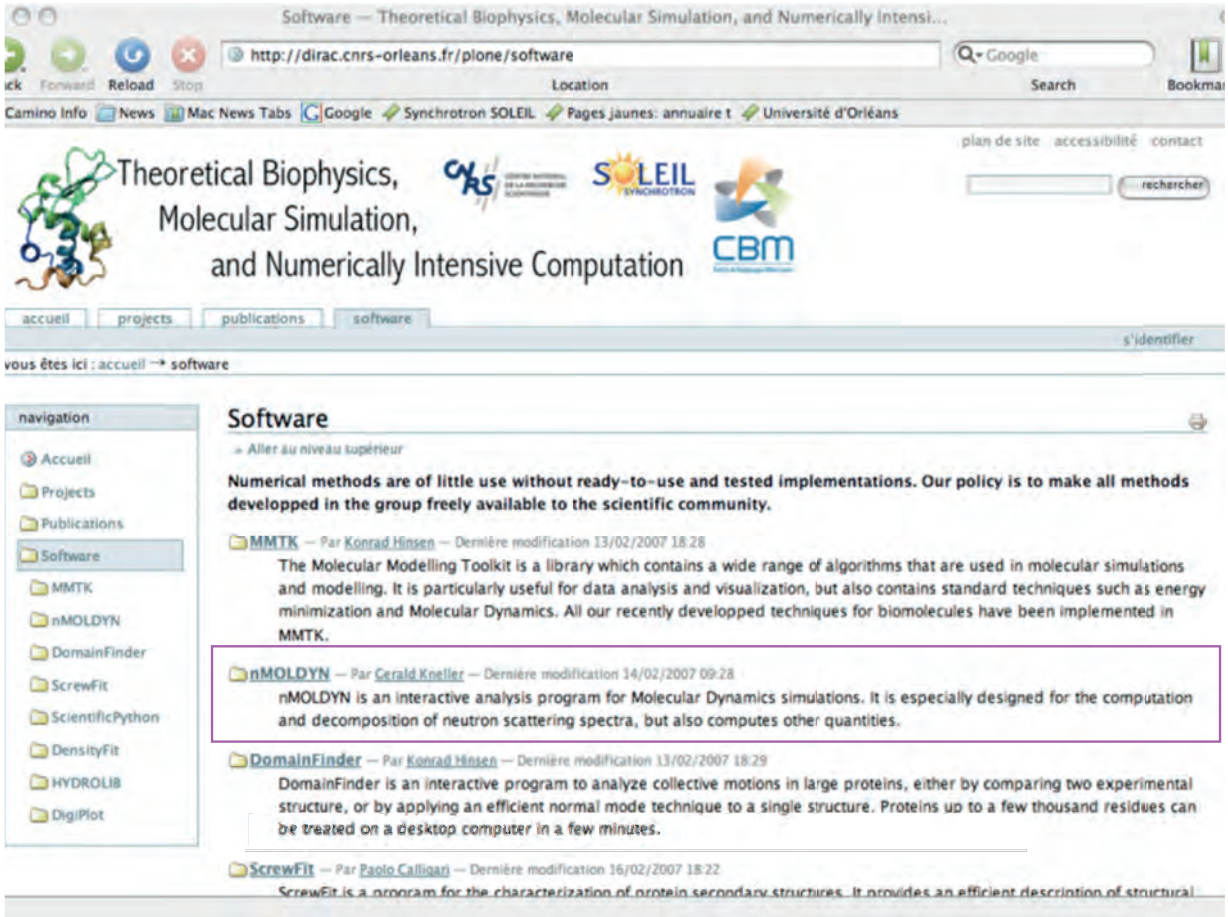
References

nMoldyn

1. G.R. Kneller, V. Keiner, M. Kneller, and M. Schiller, Comp. Phys. Commun. 91, 191 (1995).
2. T. Rog, K. Murzyn, K. Hinsen, and G.R. Kneller, J Comp. Chem. 24, 657 (2003).
3. V. Calandrini, E. Pellegrini, P. Calligari, K. Hinsen, and G.R. Kneller, Collection SFN 12, 201 (2011).
4. K. Hinsen, E. Pellegrini, S. Stachura, and G.R. Kneller, J Comp. Chem. 33, 2043 (2012).

Other

1. G. R. Kneller, Asymptotic neutron scattering laws for anomalously diffusing quantum particles, J. Chem. Phys, vol. 145, no. 4, pp. 044103–7, 2016.
2. S. Stachura and G. R. Kneller, Probing anomalous diffusion in frequency space. J .Chem. Phys., vol. 143, p. 191103, 2015.
3. G. Kneller, Quasielastic neutron scattering and relaxation processes in proteins: Analytical and simulation-based models, PCCP, vol. 7, pp. 2641 – 2655, 2005.



Merci à

- Konrad Hinsen, CBM Orléans/SOLEIL
- Mark Johnson, ILL Grenoble
- Thomas Rog, (now) U. Tampere, Finland
- Krzysztof Murzin, U. Krakow, Poland
- Eric Pellegrini, ILL Grenoble
- Gaël Goret, ILL Grenoble
- Vania Calandrini, CBM → FZJ Jülich (D)
- Paolo Calligari, CBM → SISSA (I)

ANR programme "Conception et Simulation"

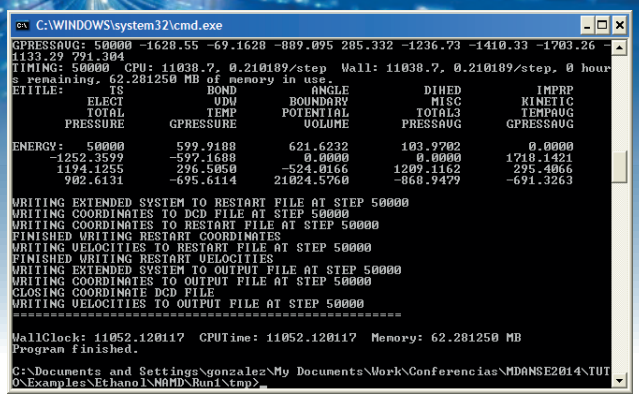
http://dirac.cnrs-orleans.fr/sputnik/home/

Analysis of MD simulations:
Validation and comparison to experiment.

Miguel A. González

(Institut Laue-Langevin, Grenoble, France)

Once our MD run is over …



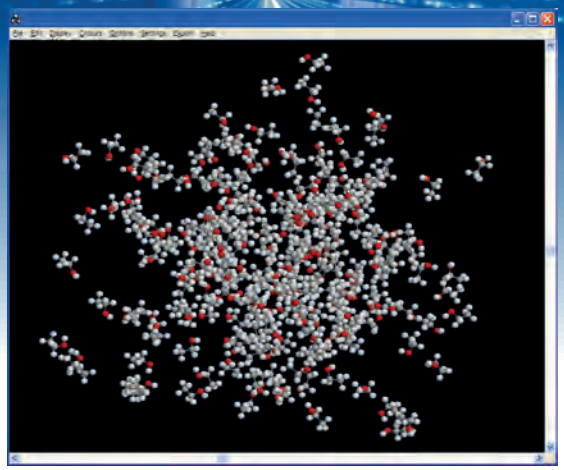
We will have a 'snapshot' of the final state of the system

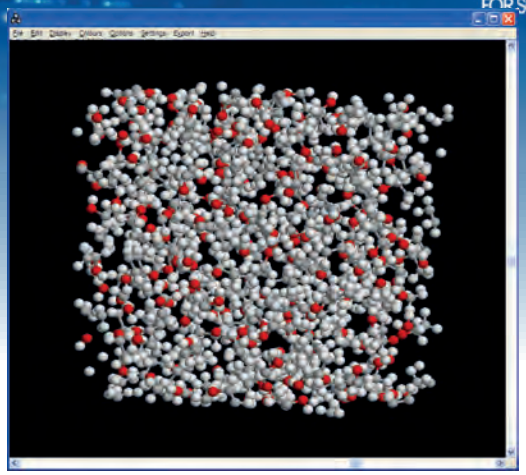
+ a trajectory file containing the positions (and velocities) at selected steps during the simulation

+ a file containing some information about the value of important quantities along the simulation, e.g. : temperature, pressure, volume, energies, etc.

The first step is to check that everything worked as expected and that we can trust the results and the trajectory given by the MD engine!

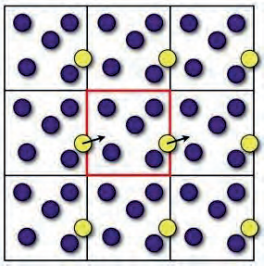
Once our MD run is over …



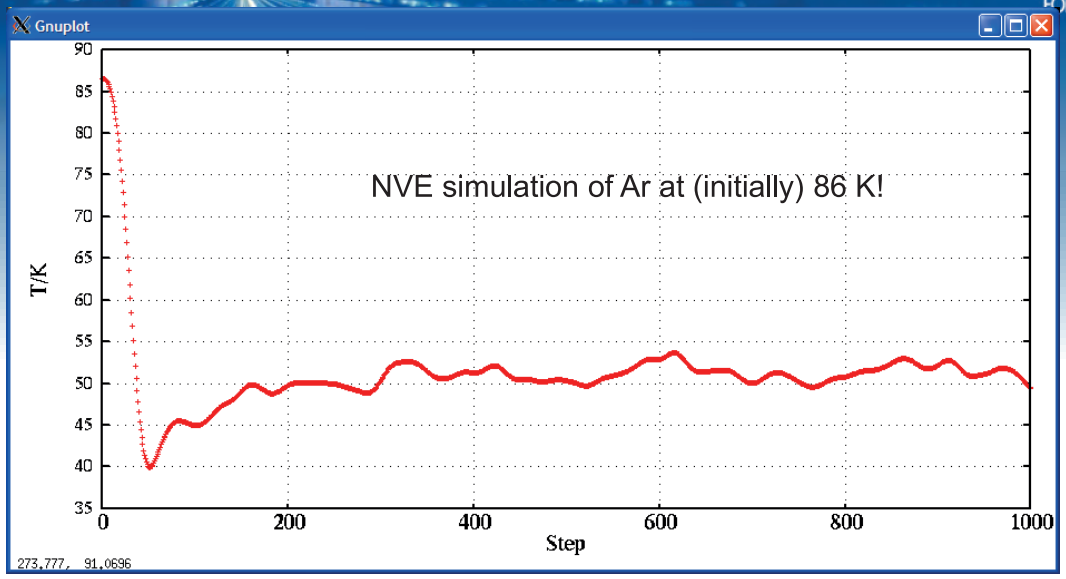


Visual checking:

- Look at final configuration and trajectory (using e.g. VMD, Rasmol, nMoldyn)
- Are molecular geometries right?
- Any 'too close' atoms?
- Don't get disturbed by PBCs



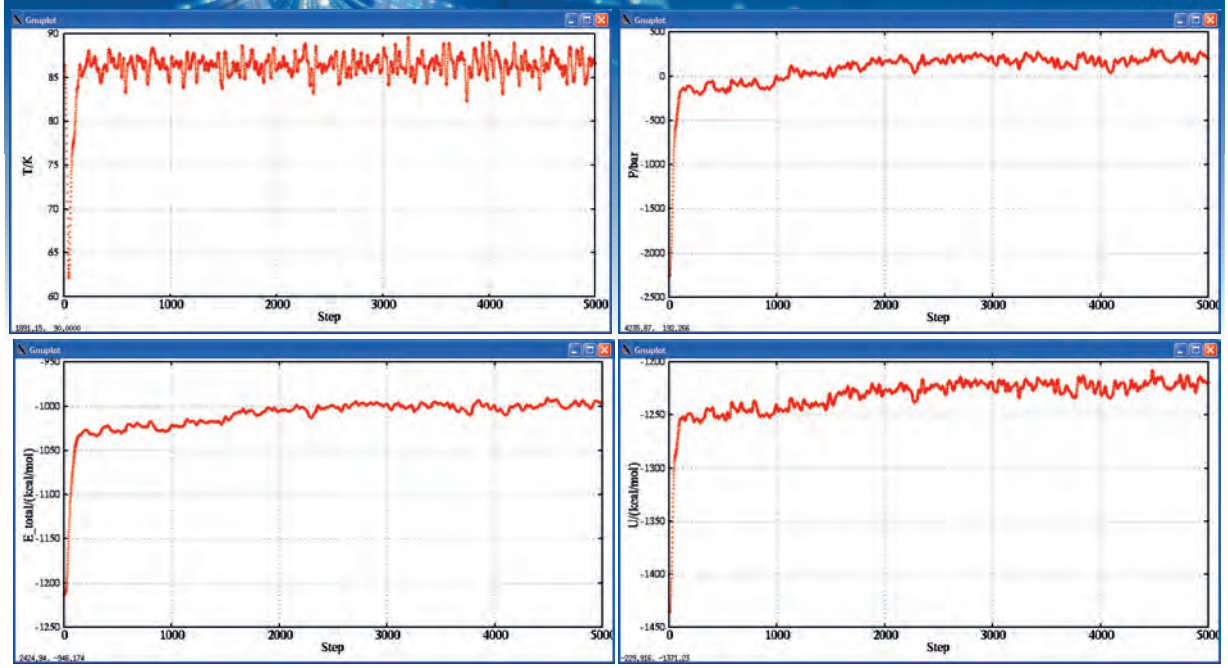
Once our MD run is over …



Check thermodynamics:

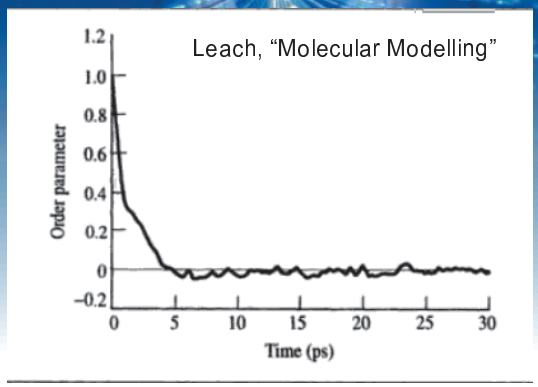
- Are the conditions (P, T, ρ) those that we want?
- Is the total energy well conserved (if NVE simulation)?
- Drifts in temperature, energy?

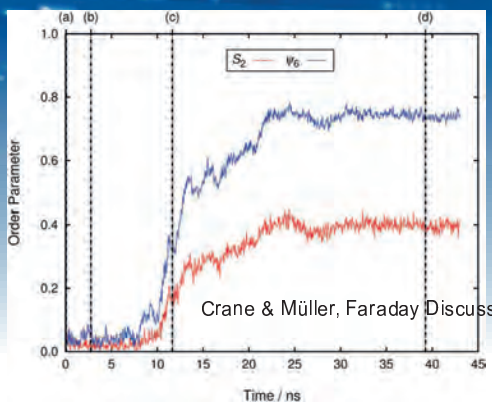
Is the system well equilibrated?



Not all the variables converge to equilibrium at the same speed!

Is the system well equilibrated?





Crane & Müller, Faraday Discuss. (2010)

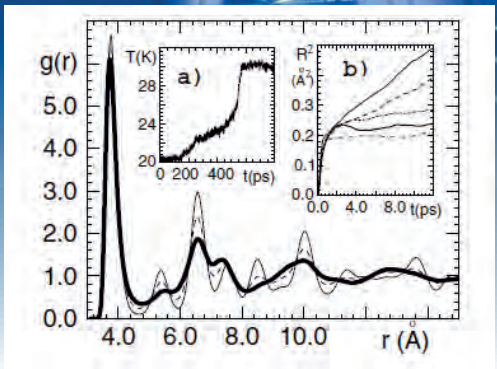
When simulating a liquid, but starting configuration is a lattice, compute evolution of appropriate order parameter to follow melting of the lattice:

$$\lambda = \frac{1}{3}\left(\lambda_x + \lambda_y + \lambda_z\right) \quad \text{with } \lambda_x = \frac{1}{N}\sum_{i=1}^{N}\cos\left(\frac{4\pi x_i}{L}\right)$$

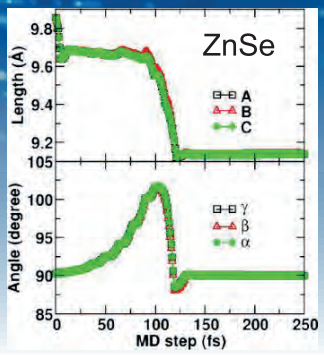
Many different order parameters: rotational order, planar order, etc.:

$$P_1 = \frac{1}{N}\sum_{i=1}^{N}\cos\gamma_i \qquad S_2 = \frac{1}{N}\sum_{i=1}^{N}3\sin^2\phi_i - 2 \qquad \psi_k = \left|\frac{1}{N}\sum_{i=1}^{N}\exp\left(ik\theta_i\right)\right|$$

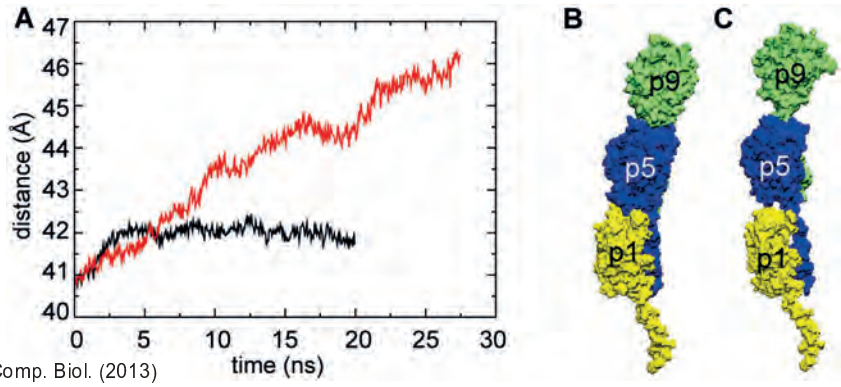
Are there changes during the simulation?



De Santis, JPCM **14** (2002)



Durandurdu, JPCM **21** (2009)



Baker, PLOS Comp. Biol. (2013)

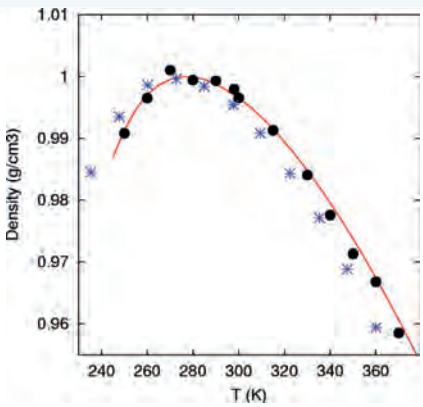
Thermodynamic averages

Apart from providing an easy way to monitor the equilibration of the simulation, thermodynamic averages serve to make a first check on the 'goodness' of our model.

MD program gives directly T, P, V, E_k, U, E_{total} and different contributions to the energy (bonding, bending, dihedral, VDW, electrostatic, etc.)

Is the density well reproduced?

How do the enthalpies of vaporization and melting compare to experiment?



$$\Delta H_{vap} = H_{gas} - H_{liq} = (U + PV)_{gas} - (U + PV)_{liq}$$

$$\Delta H_{vap} \approx \left(U_{gas}^{inter} - U_{liq}^{inter}\right) + \left(U_{gas}^{intra} - U_{liq}^{intra}\right) + \left(RT - PV_{liq}\right)$$

$$\Delta H_{vap} \approx -U_{liq}^{inter} + RT$$

Abascal & Vega, JCP (2005)

Thermodynamic properties

Other thermodynamic properties need to be computed 'a posteriori', either from the RMS fluctuations or using several simulations at different temperatures/pressures.

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V \qquad C_P = \left(\frac{\partial H}{\partial T}\right)_P \qquad \kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \qquad \alpha_P = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P$$

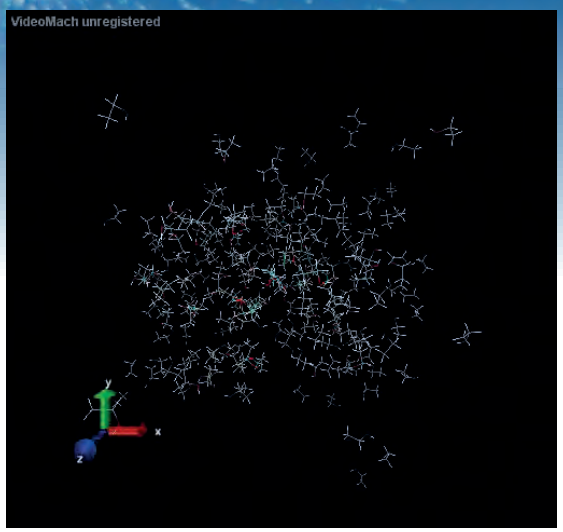
But note that the fluctuations formulae to be employed depend on the statistical ensemble used in the simulation. E.g.:

$$C_V = \frac{\left\langle \left(H - \langle H \rangle\right)^2 \right\rangle}{k_B T^2} = \frac{\left\langle \left(U - \langle U \rangle\right)^2 \right\rangle}{k_B T^2} + \frac{3}{2} N k_B T$$ (canonical ensemble, NVT)

$$C_V = \frac{9 N k_B^3 \langle T \rangle^2}{6 k_B^2 \langle T \rangle^2 - 4\left\langle \left(U - \langle U \rangle\right)^2 \right\rangle}$$ (microcanonical ensemble, NVE)

Trajectory analysis

Most properties of interest need to be calculated from the time evolution of the positions and/or velocities during the simulation.



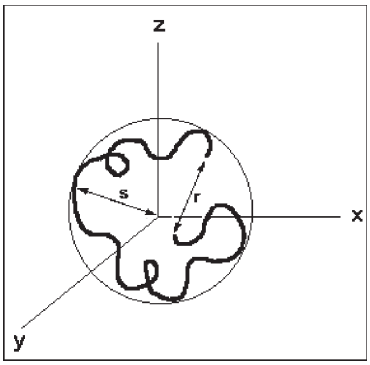
Does the system moves as expected?

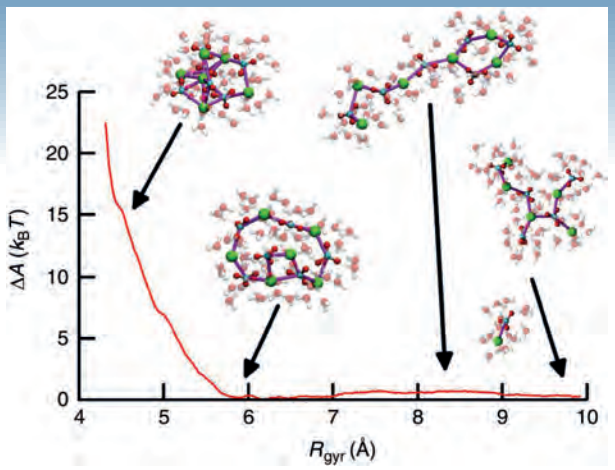
Do we observe particular motions of interest: conformational changes, CH_3 torsions, etc. ?

Radius of gyration

Useful to determine the dimensions and conformation of large molecules, e.g. polymers, macromolecules, clusters, etc.

$$R_g^2 = \frac{1}{N}\sum_{i=1}^{N}\left(r_i - r_{\text{mean}}\right)^2$$





Demichelis, Nature Comm. (2011)

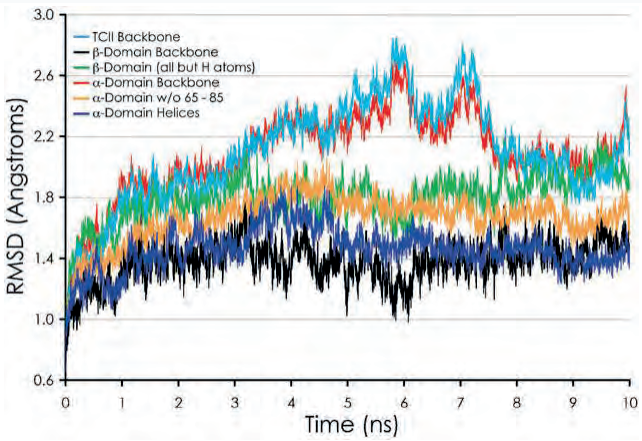
Root Mean Square Deviation (RMSD)

Useful to determine the deviation with respect to a reference structure.

Requires removing global translational and rotational motions.

Serves to determine correctness of structure or to identify conformational changes.

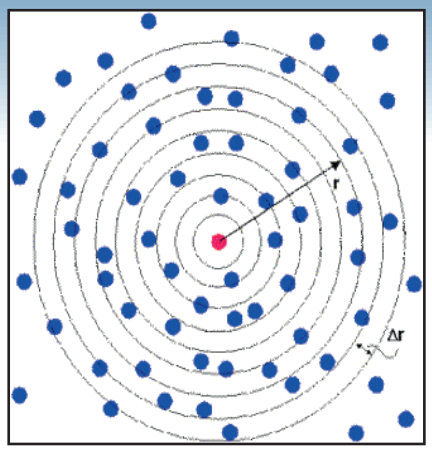
$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}\left(r_i - r_{\text{ref}}\right)^2}{N}} \quad \text{or} \quad RMSD = \sqrt{\frac{\sum_{i=1}^{N} m_i\left(r_i - r_{\text{ref}}\right)^2}{\sum_{i=1}^{N} m_i}}$$



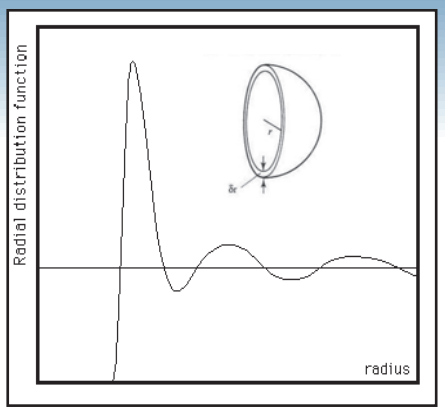
Allis, Mol. Biosyst. (2010)

Structure: Radial Distribution function

The radial distribution function (RDF) is a pair correlation function that describes how, on average, the atoms in a system are radially packed around each other.



$$g(r) = \frac{\rho(r)}{\langle \rho \rangle}$$



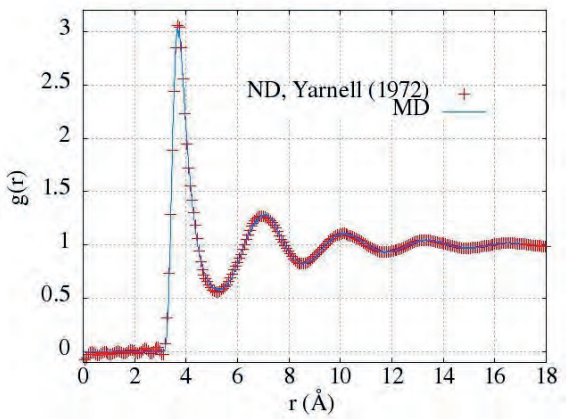
$$g(r) = \frac{V}{N^2}\left\langle \sum\sum \delta\left(r - r_{ij}\right)\right\rangle$$

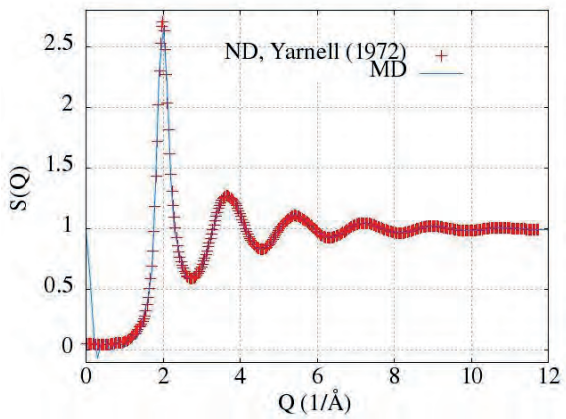
Static structure factor

Directly related to the RDF: $S(k) = 1 + 4\pi\rho\int_0^\infty r^2 \frac{\sin(kr)}{kr} g(r)dr$

Experimental observable (neutron or X-ray diffraction):

$$\rho(k) = \sum_{i=1}^{N} \exp\left(i\vec{k}\cdot\vec{r}_i\right) \qquad S(k) = \frac{1}{N}\langle \rho(k)\rho(-k)\rangle$$

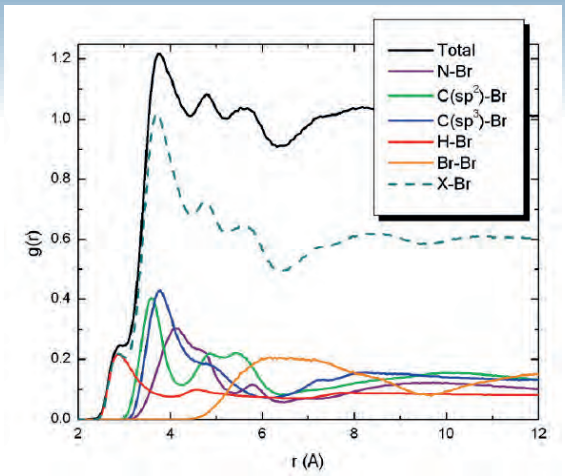


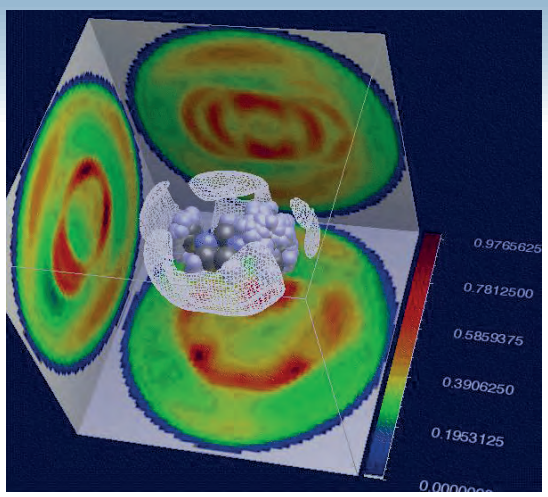


Spatial distribution functions

Possible to go beyond experiment and compute 'non measurable' quantities, such as partial r.d.f.'s, but also 3D spatial distributions.

For each molecule, transform coordinates to place the molecule at origin with a well defined orientation and fill atomic density in 3D grill.



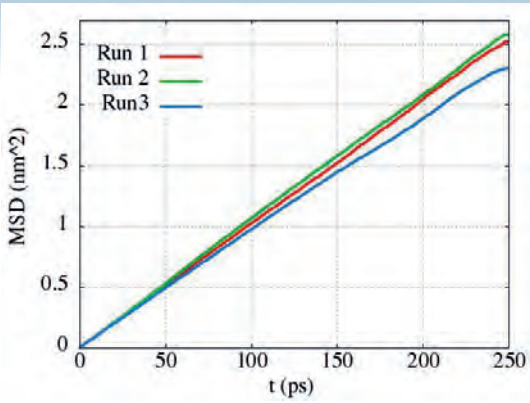


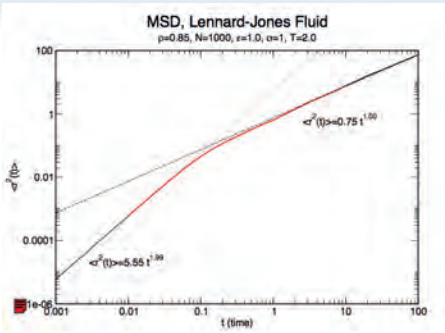
Dynamics: Mean Square Displacements

The m.s.d. is calculated in a very simple way: $\left\langle \Delta r^2(t) \right\rangle = \left\langle \left[\vec{r}(t) - \vec{r}(0) \right]^2 \right\rangle$

It provides an easy way of checking the dynamical state of the system. The self-diffusion coefficient can be obtained from Einstein's relation:

$$\left\langle \Delta r^2(t \to \infty) \right\rangle \propto 6Dt$$



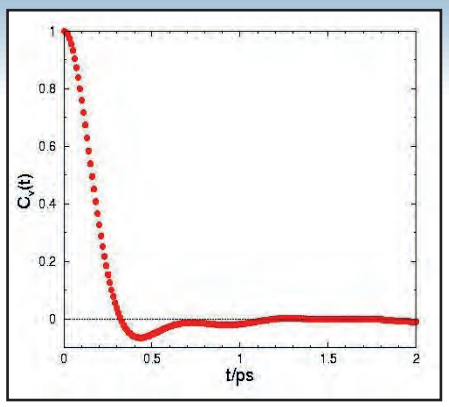


Note that when calculating the m.s.d. we average over different time origins, i.e. we compute $\langle \Delta r^2(t-t_0) \rangle$ $\Rightarrow$ errors increase with increasing t (less origins available).

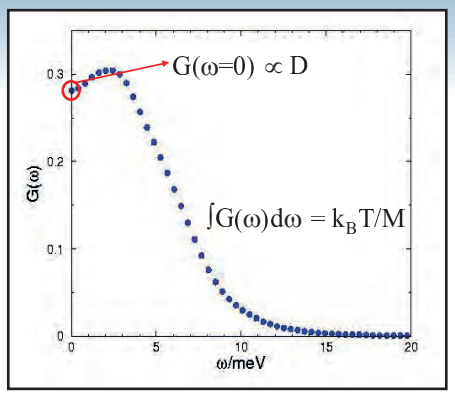
(Time) Correlation functions

$$C_{xy} = \frac{1}{N}\sum_{i=1}^{N} x_i y_i \equiv \langle xy \rangle \qquad C_{xy}(t) = \langle x(t)y(0) \rangle$$

E.g. if $x_i = v_i \Rightarrow$ Velocity Autocorrelation Function (VACF): $C_{vv}(t) = \frac{1}{N}\sum_{i=1}^{N} \vec{v}_i(t)\vec{v}_i(0)$



FT

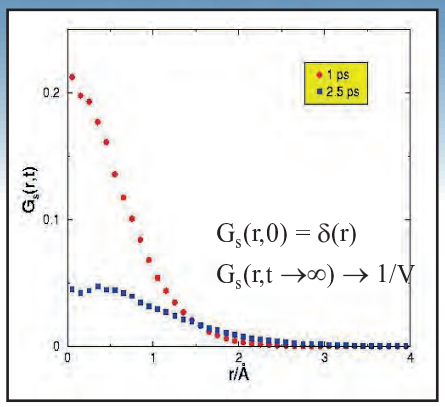


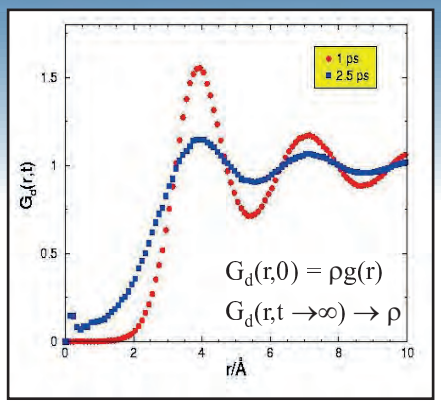
Can be used to calculate D:
$D = \int C_v(t)dt / 3$

Density of States

Van Hove correlation functions

$$G(\vec{r},t) = \frac{1}{N}\sum_i \sum_j \delta\left(\vec{r} + \vec{r}_i(0) - \vec{r}_j(t)\right) = G_{self}(\vec{r},t) + G_{distinct}(\vec{r},t)$$





The FT in space of G(r,t) gives the intermediate scattering function, F(k,t), and its FT in space and time the dynamic structure factor, S(k,ω), which can be measured with neutron scattering techniques.

Error analysis

- **Systematic errors:**
  - Numerical integration, finite-size effects, interaction cut-off.
  - They are intrinsic and reproducible.
  - For errors due to the interaction cut-off, long-range corrections can be included in some cases by assuming a uniform distribution beyond r_c.
- **Errors due to inadequate sampling of phase space:**
  - Ergodic hypothesis ($\langle\ldots\rangle_{ensemble} = \langle\ldots\rangle_{time}$) does not apply.
  - Not reproducible.
- **Statistical errors:**
  - Random fluctuations in the measurements.
  - Usual statistical analysis can be applied.

Errors in correlated measurements

If we evaluate a given property X at different times, we can use standard statistical methods to compute $\langle X\rangle$ and σ_X. But contrary to most experimental data, our X_i "measurements" can be correlated $\Rightarrow$ σ_X is underestimated.

The right variance can be computed as:

$$\sigma^2(\langle X\rangle) = \frac{1}{M}\sum_{i=1}^{M}(X_i - \langle X\rangle)^2 \times \left[1 + 2\sum_{\mu}\left(1 - \frac{\mu}{M}\right)\emptyset_\mu\right]$$

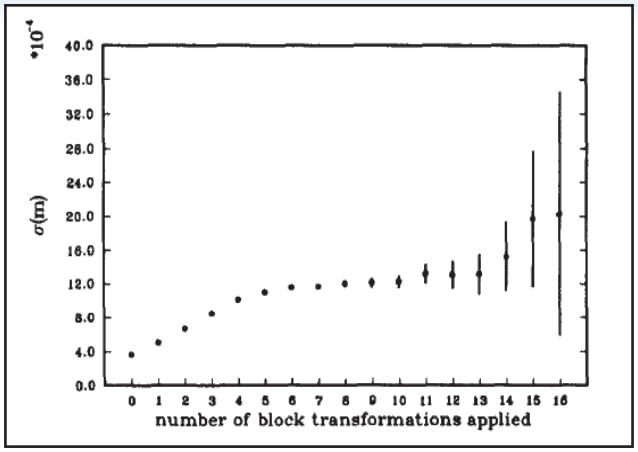
Where M is the number of measurements and ϕ_μ is the following autocorrelation function:

$$\emptyset_\mu = \frac{\langle X_\mu X_0\rangle - \langle X\rangle^2}{\langle X^2\rangle - \langle X\rangle^2}$$

Errors in correlated measurements

Block averaging method (Flyvbjert & Petersen, JCP 1989).

- Calculate averages over M_b blocks of increasing size b.
- As b↑, blocks are less correlated until becoming statistically independent.
- The minimal necessary block length and the searched estimate for the variance appear as a plateau in $\sigma^2(\langle A\rangle_b) = 1/(M_b - 1) \cdot \Sigma_\beta (A_\beta^2 - \langle A\rangle_\beta^2)$



Errors in correlation functions

$$C_X^{\mathrm{MD}}(t) = \langle X(t)X(0)\rangle_{\mathrm{MD}}$$

The variance of this autocorrelation function can be estimated as (Allen-Tildesley 1986 and Zwanzig & Ailawadi, Phys. Rev. (1969):

$$\sigma^2\left(\langle X(t)X(0)\rangle_{\mathrm{MD}}\right) \approx \frac{2\tau_X C_X(0)^2}{t_{\mathrm{MD}}}$$

And for a normalized autocorrelation function:

$$\sigma\left(\langle X(t)X(0)\rangle_{\mathrm{MD}} / \langle X^2\rangle_{\mathrm{MD}}\right) \approx \sqrt{\frac{2\tau_X}{t_{\mathrm{MD}}}} \times \left[1 - C_X^{\mathrm{MD}}(t)\right]$$

$$\frac{t_{MD}}{\tau_X} = 1000 \rightarrow \sigma \approx 0.04\ (t \rightarrow \infty)$$

$$\frac{t_{MD}}{\tau_X} = 100 \rightarrow \sigma \approx 0.14\ (t \rightarrow \infty)$$

$$\frac{t_{MD}}{\tau_X} = 10 \rightarrow \sigma \approx 0.45\ (t \rightarrow \infty)$$

But for single-particle correlation functions, the situation is much better:

$$\sigma \approx \sqrt{\frac{2\tau_X}{N t_{\mathrm{MD}}}}$$

MD analysis programs

Most MD codes contain some analysis tools and different utilities can be found in the web, but there is not a truly popular analysis software (compared to codes such as NAMD, Gromacs, LAMMPS, …).

The likely reason is that the analysis of a MD trajectory is a very "personal" task: almost an endless number of properties and selections can be done, meaning that quite often you may need to write your own program to make the analysis you want.

But nMoldyn can help you in computing all the standard properties (as those shown before) and, in particular, the neutron scattering observables: S(Q), EISF(Q), $S_{inc}(Q,\omega)$, $S_{coh}(Q,\omega)$

- First release by G. Kneller in 1995
- Completely rewritten (F77→Python) in 2003
- nMoldyn3 in 2011
- nMoldyn4 → MDANSE (beta version) 2016

MDANSE features

Read trajectory files from many different simulation programs:

- MMTK
- Amber
- CHARMM/X-PLOR
- DL_POLY
- LAMMPS
- NAMD
- Materials Studio (Discover, Forcite)
- VASP (DFT)
- CASTEP (DFT)

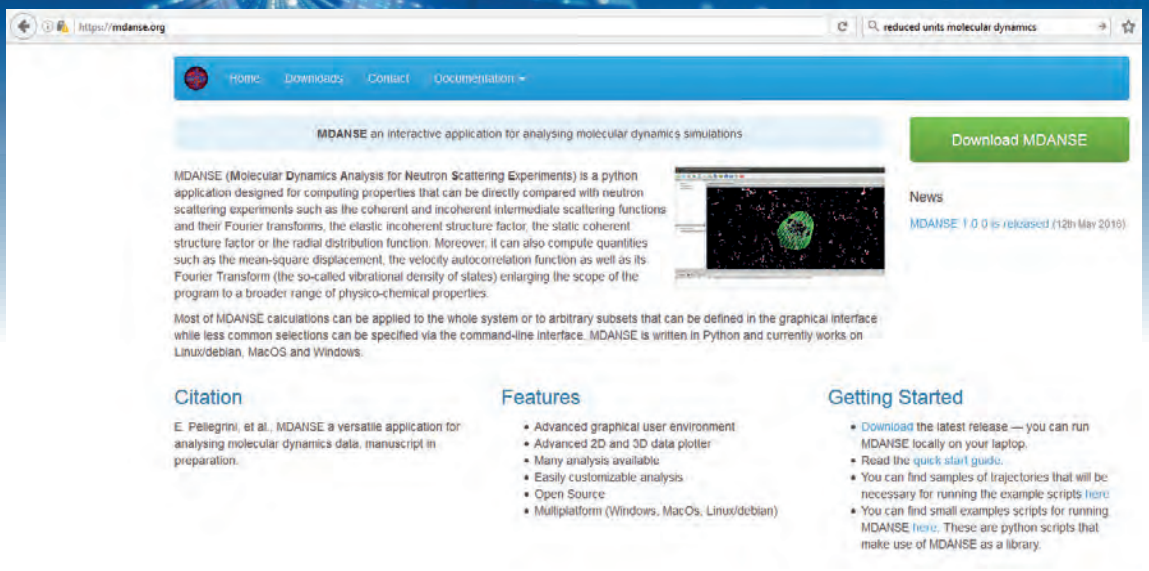
Database containing all atomic properties, including neutron scattering lengths and different isotopes, and user customizable.

Multiple possibilities to select atoms or group of atoms and to "transmutate" atoms (needed to make "isotopic substitution" analysis).

Link to McStas to perform virtual experiments

It can be run from the GUI or from a text terminal using only python scripts.

MDANSE: Hopefully soon released



Eric Pellegrini (pellegrini@ill.fr)

nMoldyn and neutron scattering observables

$$\mathcal{F}_{coh}(\mathbf{q},t) = \frac{1}{N}\sum_{\alpha,\beta} b_{\alpha,coh}\, b_{\beta,coh} \langle \exp[-i\mathbf{q}\cdot\hat{\mathbf{R}}_\alpha(0)]\exp[i\mathbf{q}\cdot\hat{\mathbf{R}}_\beta(t)]\rangle,$$

$$\mathcal{F}_{inc}(\mathbf{q},t) = \frac{1}{N}\sum_{\alpha} b^2_{\alpha,inc} \langle \exp[-i\mathbf{q}\cdot\hat{\mathbf{R}}_\alpha(0)]\exp[i\mathbf{q}\cdot\hat{\mathbf{R}}_\alpha(t)]\rangle.$$

$$\mathcal{F}_{\rm coh}(\mathbf{q},t) = \sum_{I,J\geq I}^{N_{species}} \sqrt{n_I n_J \omega_{I,\rm coh}\omega_{J,\rm coh}}\,\mathcal{F}_{IJ,\rm coh}(\mathbf{q},t),$$

$$\mathcal{F}_{\rm inc}(\mathbf{q},t) = \sum_{I=1}^{N_{species}} n_I \omega_{I,\rm inc}\mathcal{F}_{I,\rm inc}(\mathbf{q},t)$$

$$EISF(\mathbf{q}) \doteq \lim_{t\to\infty}\mathcal{F}_{\rm inc}(\mathbf{q},t).$$

$$\mathcal{F}_{IJ,\rm coh}(\mathbf{q},t) = \frac{1}{\sqrt{n_I n_J}}\sum_{\alpha}^{n_I}\sum_{\beta}^{n_J}\langle \exp[-i\mathbf{q}\cdot\hat{\mathbf{R}}_\alpha(t_0)]\exp[i\mathbf{q}\cdot\hat{\mathbf{R}}_\beta(t_0+t)]\rangle_{t_0},$$

$$S_{\rm coh}(q) = \sum_{I,J\geq I}^{N_{species}} \sqrt{n_I n_J \omega_{I,\rm coh}\omega_{J,\rm coh}}\,S_{IJ}(q)$$

$$\mathcal{F}_{I,\rm inc}(\mathbf{q},t) = \frac{1}{n_I}\sum_{\alpha=1}^{n_I}\langle \exp[-i\mathbf{q}\cdot\hat{\mathbf{R}}_\alpha(t_0)]\exp[i\mathbf{q}\cdot\hat{\mathbf{R}}_\alpha(t_0+t)]\rangle_{t_0}.$$

So nMoldyn is extremely useful in computing the same neutron scattering functions that we can measure in e.g. a time-of-flight or a backscattering spectrometer, but ...

Comparing simulation and experiment

You still need to think carefully which are the exact quantities that we are measuring and computing.

For example, for the EISF (Elastic Incoherent Structure Factor):

The EISFs were computed from the MD trajectories with the relation:

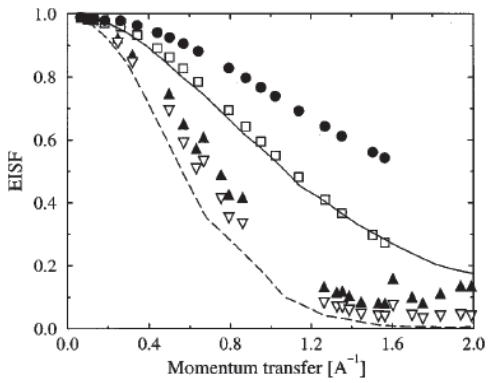
$$\mathrm{EISF}(Q)=\sum_j b_j^2 |\langle\langle e^{-i\mathbf{Q}\cdot\mathbf{R}_j(t)}\rangle_t\rangle_{\hat{Q}}|^2, \quad (19)$$

$$EISF(\mathbf{q}) = \lim_{t\to\infty} \mathcal{F}_{\mathrm{inc}}(\mathbf{q}, t),$$

In the quasielastic energy transfer range, the dynamical structure factor can be written as:[9]

$$S(\mathbf{Q},\omega)=e^{-Q^2\langle u^2\rangle/6}\times[I^D(\mathbf{Q},+\infty)\delta(\omega)+\Gamma(\mathbf{Q},\omega)] + Background, \quad (14)$$

$$\mathrm{EISF}(\mathbf{Q})=I^D(\mathbf{Q},+\infty).$$



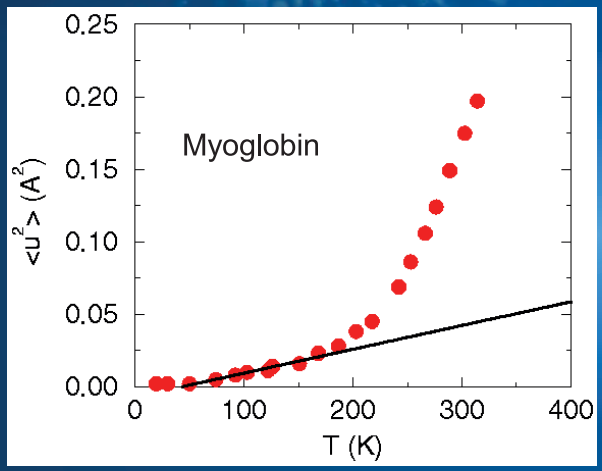
Due to the experimental EISF extraction procedure, the vibrational librational motions do not contribute to the measured EISF. In constrast, the simulation-derived EISFs obtained from Eq. (19) are computed without decomposition of the atomic trajectories into their vibrational and diffusive parts. The relation between the experimental and simulated elastic incoherent structure factors is therefore, with a spherical averaging:

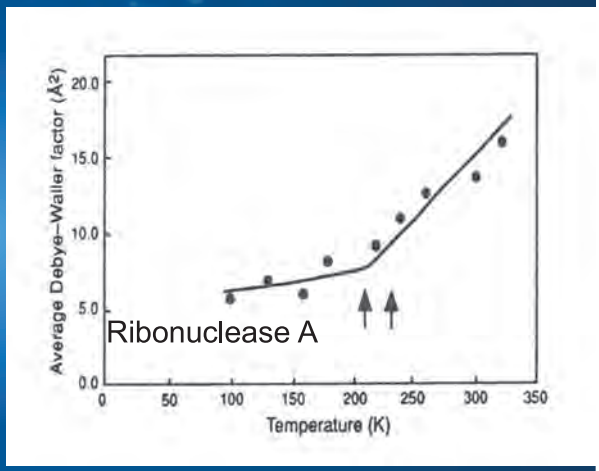
$$\mathrm{EISF}_{\mathrm{sim}}(Q)=e^{-Q^2\langle u^2\rangle/6}\,\mathrm{EISF}_{\mathrm{exp}}(Q). \quad (20)$$

From Morelon, J. Chem. Phys. (1998)

Comparing simulation and experiment

$\langle u^2\rangle$ and dynamical transition in proteins from elastic neutron scattering experiments:



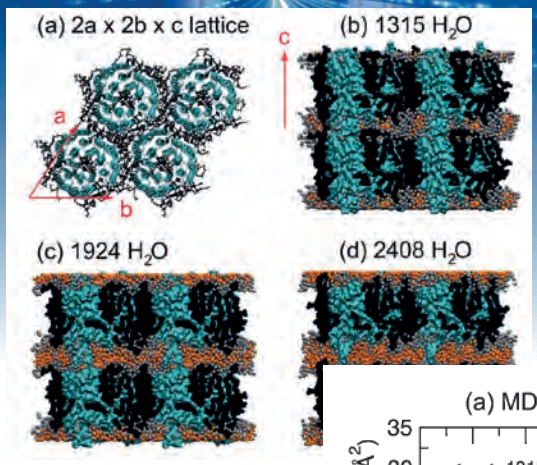


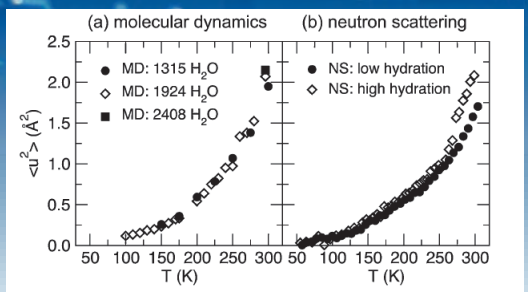
Doster *et al*, Nature (1989)

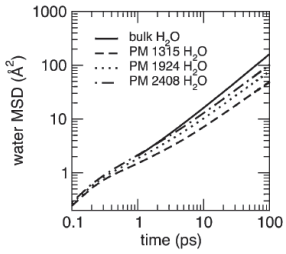
Rasmussen *et al*, Nature (1992).

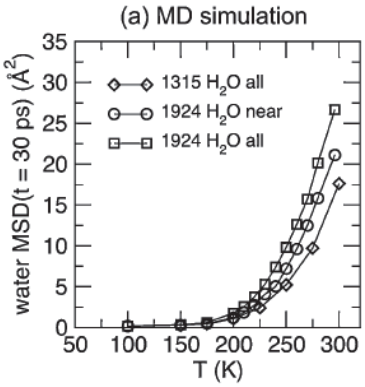
Dynamic transition in $\langle u^2\rangle$ correlates with the onset of enzymatic activity!

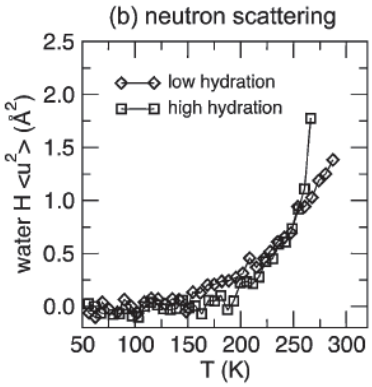
Comparing simulation and experiment







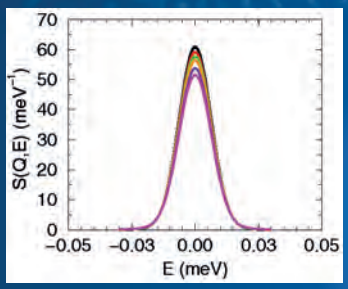


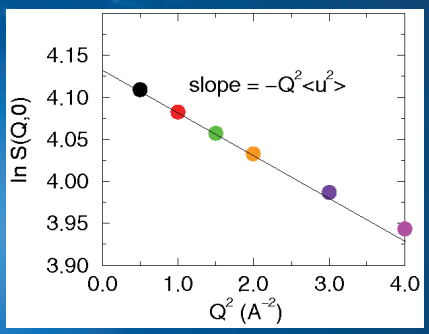


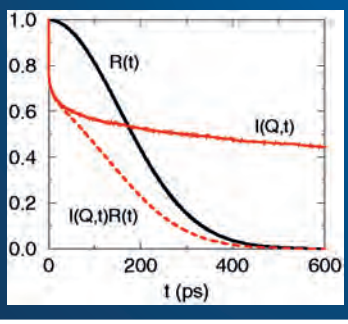
From Tobias, Faraday Discuss. (2009)

Computing $<u^2>$ as in experiment

Experiment: estimate mean-squared displacement from elastic intensity via Debye-Waller factor: $I(0) = \exp(-Q^2<u^2>)$







Simulation: calculate resolution-broadened S(Q,E) as FT of I(Q,t)R(t), where R(t) is the FT of the instrument resolution function

$$I(Q,t) = \frac{1}{N}\sum_j e^{i\vec{Q}\cdot\vec{r}_j(t)} e^{i\vec{Q}\cdot\vec{r}_j(0)}$$

Comparing simulation and experiment

K.[45,47] Over the temperature range (>250 K) where the water content is different, the neutron data show that water mobility at a given temperature is higher in the system that contains more water, in agreement with the simulations. Second, the simulations predict that the extent of water motion on a time scale of tens of ps is roughly an order of magnitude greater than the fluctuations measured by neutron scattering on the ns timescale above 200 K. This discrepancy is largely due to the differences in the way the quantities are measured. The simulation data are computed from trajectories in real space, while the neutron data are obtained by fitting the Q-dependence of the intensity of the resolution-broadened elastic scattering, I_{el}, to the Debye–Waller form, $I_{el}(Q) = A\exp(-Q^2\langle u^2\rangle/6)$, over a restricted range of Q (*e.g.* $Q^2 < 0.9$ Å^2).[45,47] To check for consistency with the neutron data, we have computed the resolution broadened dynamic structure factors, $S(Q,E)$, for water H atoms in the system with 1924 water molecules at 296 K as described elsewhere,[64,72] using the energy resolution of the neutron spectrometer on which the measurements reported in Fig. 8b were made. Fitting the elastic intensities, $I_{el} = S(Q,0)$, obtained at three values of Q^2 between 0.30 and 0.90 Å^2 to the Debye–Waller form, we obtain a value of $\langle u^2\rangle = 2.4$ Å^2, which is in reasonable agreement with the neutron data, and much lower than the ~20 Å^2 computed directly from the trajectory.

From Tobias, Faraday Discuss. (2009)

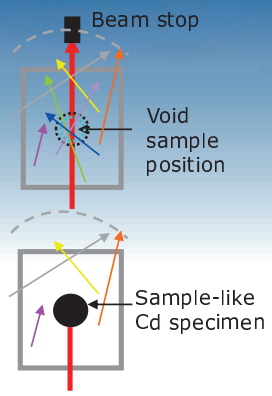
But even comparing directly computed and measured S(Q,ω) can be tricky, because is not always straightforward to get the true S(Q,ω) ...

Instrumental effects: Background

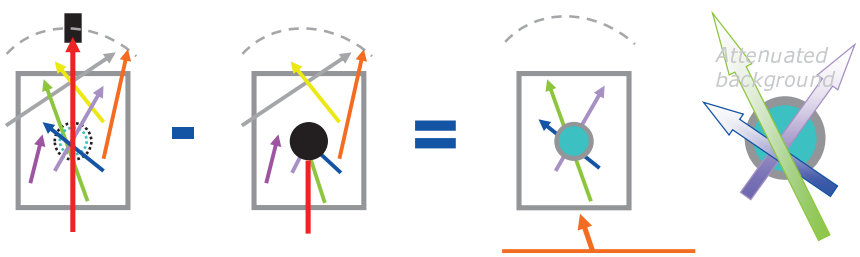
The background is sample dependent, so we need to perform two measurements:

1) *An empty beam run (container without sample), $I_{EB}(\theta,\omega)$:* Approximate measure of instrumental noise, but true background in sample measurement will be less (because attenuation due to sample scattering and absorption).

2) *A run with a "full" absorber (Cadmium or $^{10}B_4C$) in place of the sample+container system, $I_{Cd}(\theta,\omega)$:* Measure electronic noise + other sources of noise not affected by the presence of sample + container.



The background to subtract to the sample measurement can be estimated as:



$$I_b^{(s+c)} \approx I_{Cd} + T_{sc}\,(I_{EB} - I_{Cd})$$

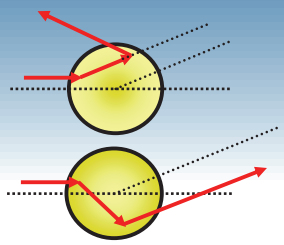
Multiple Scattering

If a neutron is scattered once, it can be scattered again ...

Multiple scattering affects the expected signal (i.e. the single-scattering intensity $I(1)$, related to the double-differential cross section) in a two-fold way:

It removes singly scattered neutrons from the original direction. Therefore it ATTENUATES the *single scattering intensity* detectable at a given angle

It contributes to the intensity detected at '*another*' angle. Therefore it INTENSIFIES the signal, mixing up with the *true single scattering* component at a given angle.



Multiple scattering causes both the loss of "good" neutrons and the detection of "bad" ones!

$$I^{\exp} \approx I^{(1)\exp} + I^{(m)\exp} \quad \text{with} \quad I^{(m)\exp} \approx I^{(2)\exp} + \ldots$$

Big problem: The evaluation of MS requires to know exactly what the neutron does in the sample, i.e. S(Q,ω), which is not known!

Some limitations of MD

Force Field Accuracy:
The results can only be as good as the underlying force field.

No explicit treatment of electrons:
Only nuclear positions available.
No chemical reactions.

Classical dynamics (no quantum effects):
Unsuitable for systems/states where quantum effects are important.
If $h\nu > k_B T$, ask yourself about quantum effects.

E.g. $k_B T$ (T=298K) = 207 cm^{-1} ≈ 6×10^{12} Hz ≈ 26 meV

Typical stretching frequencies ~ 1000 – 3800 cm^{-1}

Typical bending frequencies ~ 600 – 1650 cm^{-1}

But t ~ 1/ ν ~ 0.2 ps, so modes with characteristic time scales in the ps range or longer are reasonably treated by classical physics.

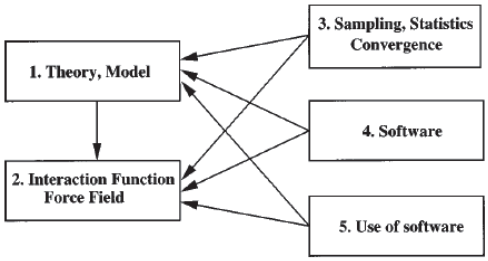
Validation of MD results

W. F. van Gunsteren and A. E. Mark,
J. Chem. Phys. **108**, 6109 (1998)

And when we have done and analyzed our simulation, and found a new result, how can we validate it?

There are at least 5 points to consider:

1. Quality of the model
2. Accuracy of the force field
3. Degree of sampling, statistics and convergence
4. Quality of the software
5. Competence in using the soft



Validation of MD results

Normally validated by comparing to experiment, but ...

- Simulation and experiment can AGREE because
  1. Simulation reflects adequately our system
  2. The property compared is not too sensitive to simulation details
  3. Compensation of errors.
- Or simulation and experiment may NOT AGREE because
  1. The simulation does not reflect well our system (due to problems in any of the 5 points mentioned above).
  2. The experimental data are not correct

Validation of MD results

Quality of the model

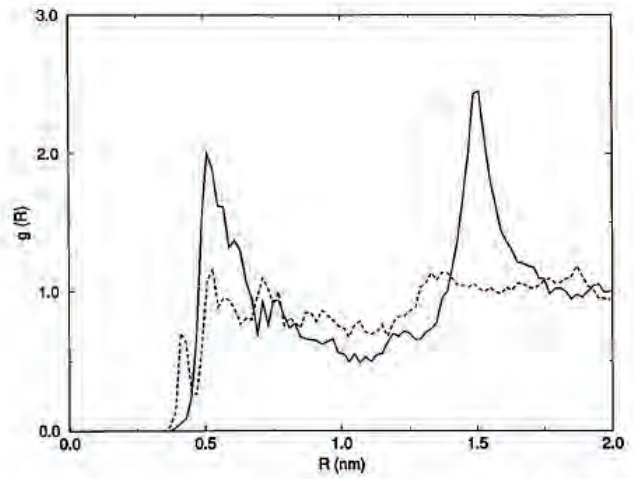
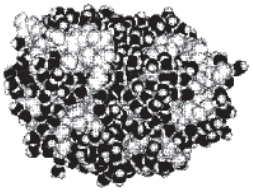


FIG. 3. Radial distribution function $g(R)$ between chlorine ions calculated from MD simulations of a periodic box containing 40 Na^+, 40 Cl^-, and 2127 water molecules using different treatments of long-range electrostatic forces; spherical truncation at $R = 1.5$ nm (solid line) and using a reaction field (dashed line). Data taken from Ref. 29.

Validation of MD results

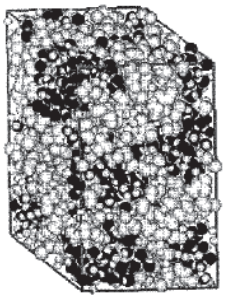
Quality of the model

Tarek & Tobias, JACS (1999)



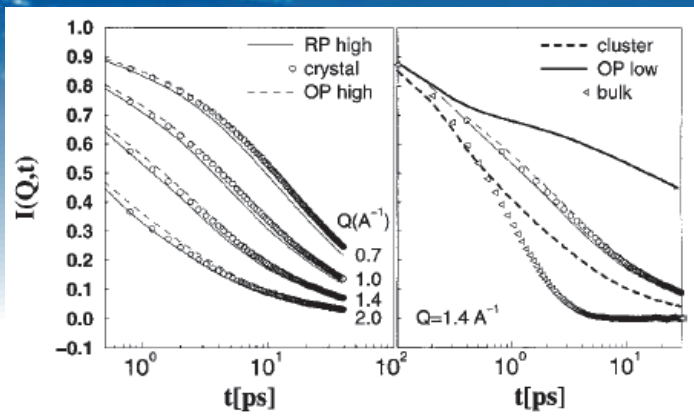
(a)

1 protein +
399 water
h = 0.58



(b)

2 proteins +
817 water
h = 0.59



This suggests that the neutron experiments performed on powders with $h \approx 0.4$ are indeed relevant to describe the water dynamics near proteins in nativelike environments such as crystals. In contrast, the cluster result shows that the water moves much too freely on the protein surface and demonstrates that cluster simulations are not appropriate for quantitative characterization of the dynamics of protein hydration water. In the low hydration

Validation of MD results

Accuracy of the force field

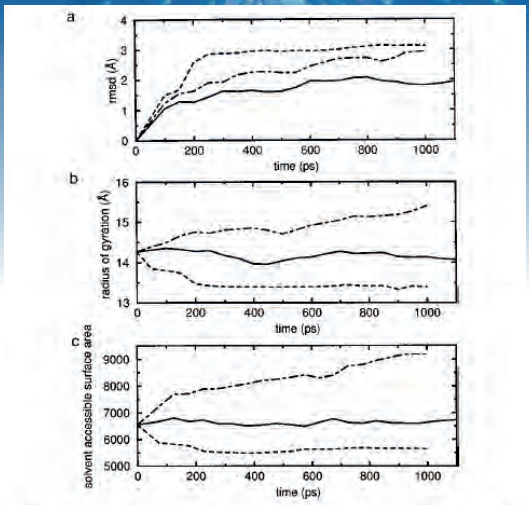


FIG. 4. Structural properties of hen egg white lysozyme (HEWL) as a function of MD simulation time using different force fields; vacuum simulation and corresponding GROMOS force field (dashed line), simulation in water using the GROMOS87 force field (dot–dashed line) and using the GROMOS96 force field (solid line); (a) root-mean-square positional difference for the C_α atoms from the crystal starting structure (Å), (b) radius of gyration (Å), (c) solvent accessible surface area (Å). Data taken from Ref. 39.

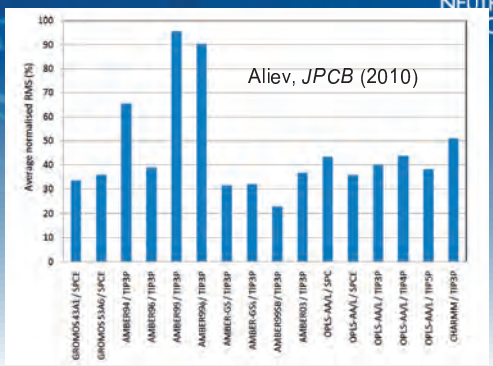
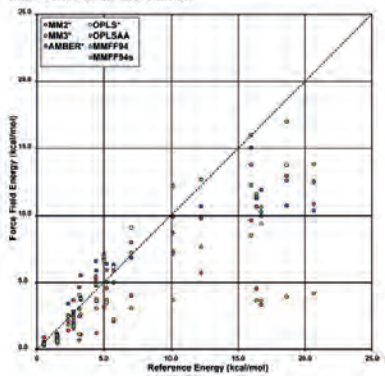


Chart 1. Comparison of Force Field Interaction Energy with Ab Initio Values for the S22 Data Set[a]

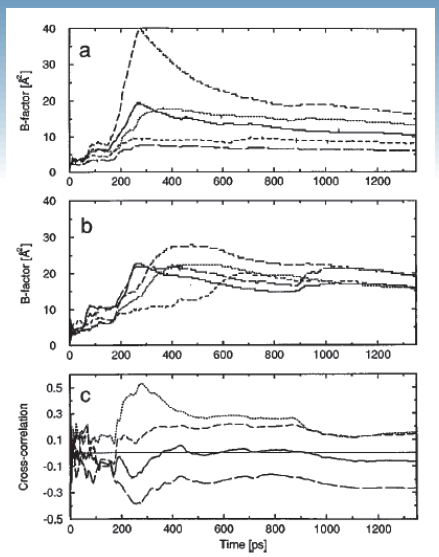


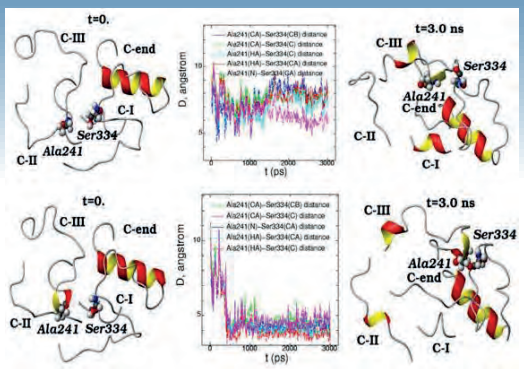
Paton & Goodman, *J. Chem. Inf. Model* (2009)

Validation of MD results

Degree of sampling, statistics and convergence

We need to ensure that $t_{equil} > \tau_{relax}(Q)$ and $t_{sample} >> \tau_{relax}(Q)$, where Q is the property we are interested in.

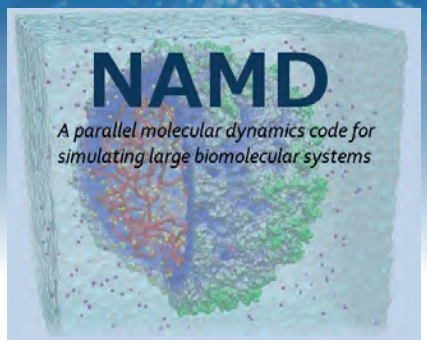


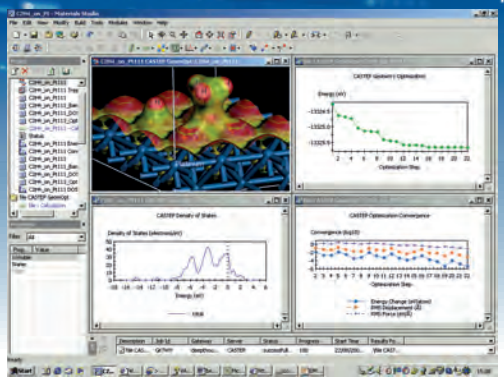


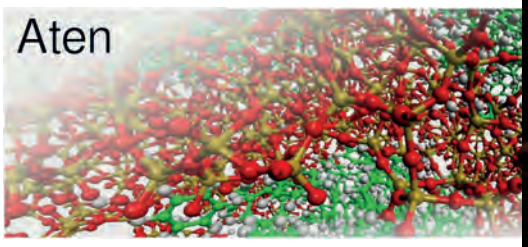
But also that we sample through all the relevant states or conformations.

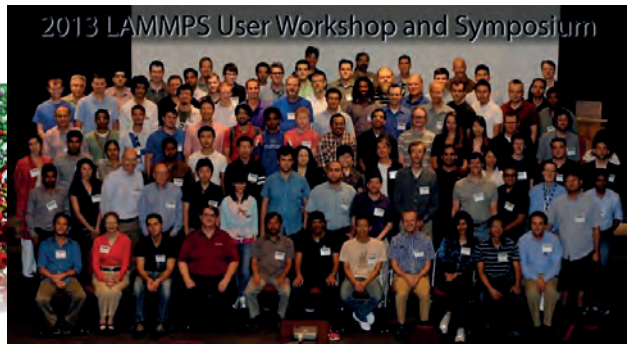
Validation of MD results

Quality of the software







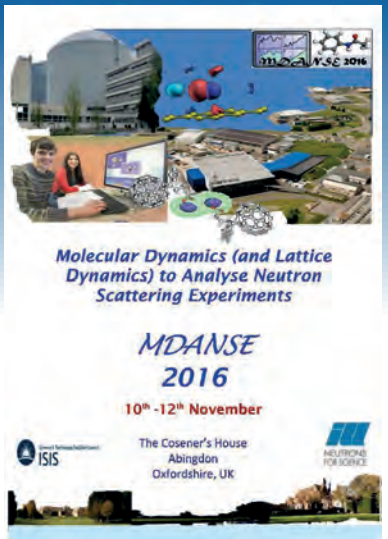


Validation of MD results

Competence in using the soft

Validation of MD results

Competence in using the soft



Thanks!

Please suggest ideas (forum, format of school, etc.) and ways to improve.

THE UNIVERSITY of York

Path Integral Molecular Dynamics

Matt Probert
Condensed Matter Dynamics Group
Department of Physics,
University of York, U.K.
http://www-users.york.ac.uk/~mijp1

Overview of lecture

- Background Theory
  - Statistical mechanics
  - Action
  - Quantum propagators
  - Path integral QM
- PIMD in CASTEP
- CASTEP PIMD examples

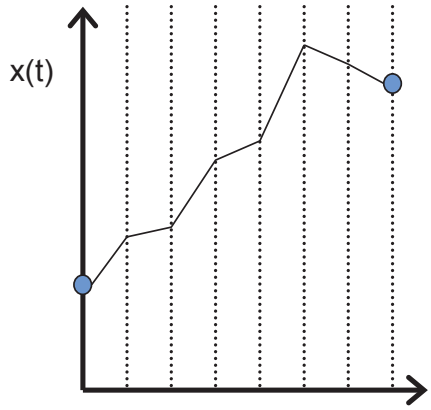
Background Theory

What is PIMD?

- Path-Integral
  - We use the Feynmann Path-Integral formulation of quantum mechanics to incorporate an approximate quantum treatment for the ions independent of how the electrons are treated
- Molecular Dynamics
  - We can use either classical or *ab initio* methods to generate forces & energies and then add in extra contributions from PI and hence move the atoms to generate quantum dynamics (PIMD)
  - This motion is fictitious and does NOT represent the real dynamics of the system
  - BUT ensemble averages of the PIMD are equivalent to the QM expectation values at an appropriate temperature
  - Hence can use PIMD to incorporate the effects of finite temperature and QM properties of the nucleus into our calculation
  - Hence include effects of zero-point motion, tunnelling, etc.

Path Integral view of QM

- The probability amplitude for a particle beings at some (x’,t’) is given by the probability of it coming from some starting point (x,t) and then summing over all possible starting points!



A possible path through space-time. The dotted lines indicate possible positions that a path could pass through at each time slice. The propagator integrates over all such possible positions, keeping the end points fixed.

The Path Integral

- Integration over all possible paths is done by *time slicing*, i.e. discretizing the path into a number of slices in time, performing the space integration at each slice, and then letting the number of slices go to infinity.
- Mathematically:

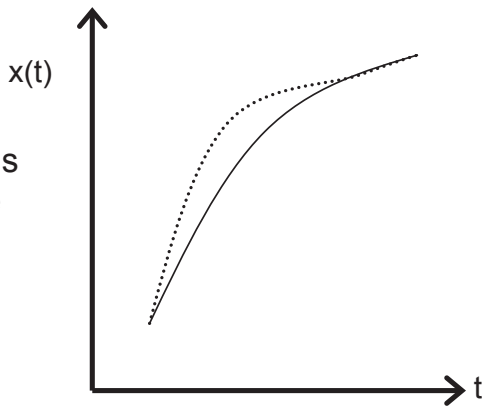
$$\int Dx \equiv \lim_{N\to\infty,\Delta t\to\infty} \int dx_{N-1} \int dx_{N-2} \cdots \int dx_1$$

Action and Classical Paths

- The *action* is defined as

$$S\left[x(t)\right] = \int_{t1}^{t2} dt\, L\left[x(t)\right]$$

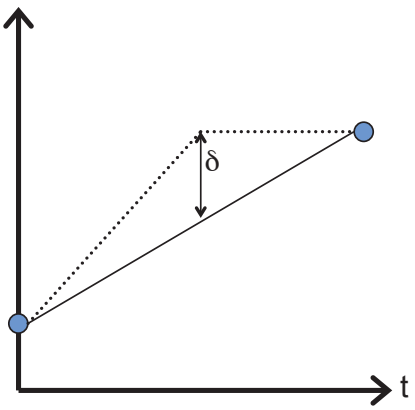
  - where the Lagrangian is L=T-V
- The classical path is that for which the action is a *minimum*



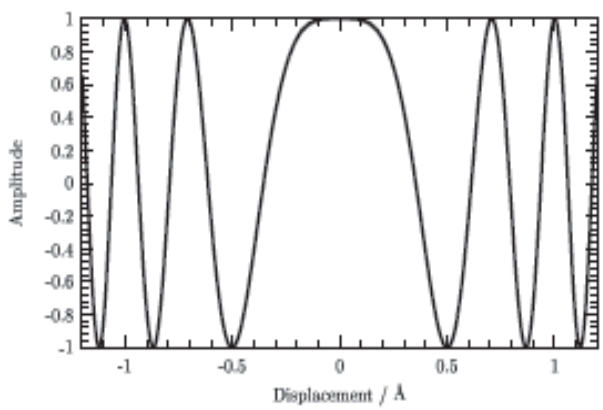
A ball moving under gravity. The solid line is the classical path, whereas the broken line is a *close* path which has a greater action. The classical trajectory balances potential and kinetic energy to minimise the overall action.

Phase Factors

- The action causes a complex phase factor which causes interference when adding neighbouring paths, e.g. H atom:



The deviation between two possible paths at a single time slice.



Classical-Quantum Crossover

- Oscillations very rapid for large path differences δ
  - For a classical 1 kg mass we see same shape shape as figure 4 but oscillation over ym scale!
- A system becomes quantum when action S~ħ
  - Oscillations add constructively over width of central peak, and then decohere
- Temperature and mass also matter
  - see quantum effects when size ~ thermal wavelength $\lambda = \sqrt{\frac{2\pi\beta\hbar^2}{m}}$

Link to Classical Stat Mech

- It can be shown that QM in imaginary time with Path Integrals is equivalent to classical statistical mechanics at finite temperature!
- If we want the properties of a particle at some (x,t) then the paths in the path integral begin and end on the same point which means can do momentum integral analytically and imaginary time is cyclic.
- In practice we discretise the path integral into *P* slices and converge w.r.t. *P*

Discretising

- Skipping details, we finally arrive at:

$$Z = Tr\left(\rho\right) = Tr\left(\exp\left(-\beta H\right)\right) = \lim_{P\to\infty} Tr\left(\exp\left(-\beta H / P\right)\right)^P$$

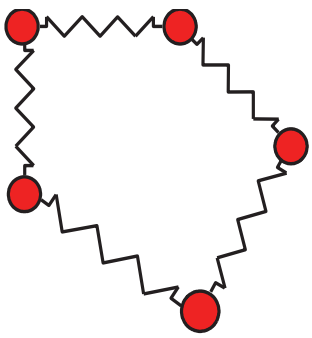
$$Z_P \sim \int dx_1 \ldots dx_P \exp\left(-\beta \sum_{s=1}^{P}\left(\frac{mP}{2\beta^2\hbar^2}\left(x_{s+1} - x_s\right)^2 + \frac{V\left(x_s\right)}{P}\right)\right)$$

where x_s is the position of at one value of the time slice in imaginary time, and $x_{P+1}=x_1$ due to cyclic nature.

- Hence the 'beads on springs' model :

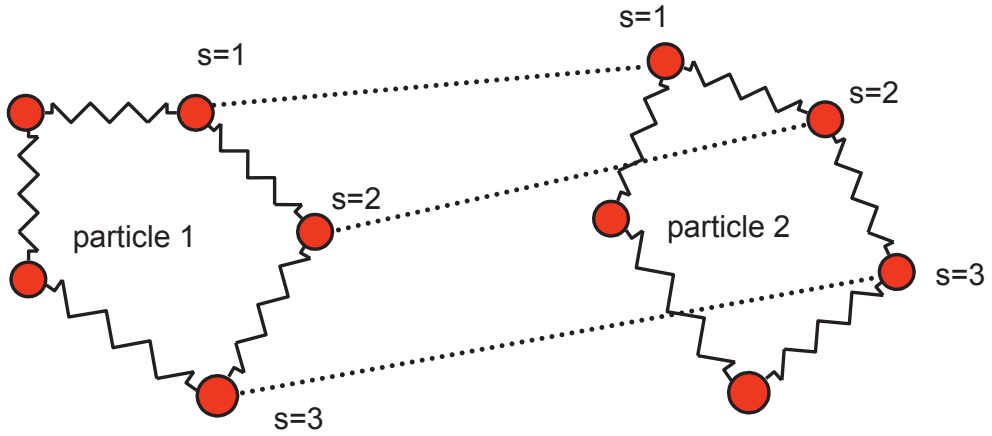
Beads on Springs

- Discretised Path Integral
- SHO interaction between nearest neighbours in imaginary time
- 1/P reduction in effect of potential
- Spring constant $k=mP/\beta^2\hbar^2$
- Hence springs get stiffer at high T
  - classical limit of a single bead
- Floppy springs at low T
  - QM delocalisation
- Centroid corresponds to classical position



Path integral view of a single quantum particle.

Multiple QM Particles



- Spring interaction only within a single particle
- Conventional V/P interaction at equivalent values of imaginary time between particles

Harmonic Issues

- The spring interaction has a fundamental frequency + harmonic modes
  - Need to integrate these accurately with MD to get proper ensemble distribution
  - Ergodicity problems => cannot use NVE or simple Nose-Hoover thermostat
  - Use N-H chain or Langevin
  - In CASTEP PIMD can only use Langevin at moment
- Also, k~P so frequency increases as converge number of beads
  - so must *reduce* MD timestep => more expensive
- Unless use *staging modes* or *normal modes*
  - Transform bead masses to compress the intra-bead spectrum and hence keep timestep constant as increase P

More

- It is possible to go beyond simple cyclic paths to consider open paths
  - Hence can sample the momentum distribution – relevant to Vesuvio experiments
  - Above uses closed paths and hence do the momentum integral analytically
- Also possible to go beyond ensemble averages and consider time evolution
  - Centroid PIMD, Ring polymer MD, or partially-adiabatic PIMD

PIMD in CASTEP

CASTEP keywords

- Usual SCF & MD keywords PLUS

  `md_use_pathint=true`

  `md_num_beads=16`

  `num_farms=16`

  `md_pathint_staging=true`

  `md_num_stages=1`

- Restrictions
  - **`num_farms=1`** or **`md_num_beads`**
  - no constraints
  - only Langevin thermostat

Analysis

- Materials Studio does not support PIMD
- The `.castep` file gives a brief summary of what is happening in the user units …

```
          ======> Path integral bead no. 003 <=====

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
x                                                MD Data:          x
x                                                                  x
x              time :       0.001000                   ps         x
x                                                                  x
x   Potential Energy:    -543.432706                   eV         x
x   Kinetic   Energy:       0.034494                   eV         x
x   Total     Energy:    -543.398212                   eV         x
x   Hamilt    Energy:    -543.397578                   eV         x
x                                                                  x
x        Temperature:     266.854751                    K         x
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

Analysis

- More advanced analysis requires more data, for which we use the `.md` file.
- This contains a LOT of information, for each time step, always using atomic units:

```
                          1.19476569E+004
                -1.99707968E+001     -1.99692125E+001      9.64993404E-004  <-- E
                  6.43328936E-04                                            <-- T
                 1.32280829E+001      0.00000000E+000      0.00000000E+000  <-- h
                 0.00000000E+000      1.32280829E+001      0.00000000E+000  <-- h
                 0.00000000E+000      0.00000000E+000      1.32280829E+001  <-- h
N      1         4.83250673E+000      3.95868000E+000     -3.95873877E+000  <-- R
N      2         4.61612393E+000      5.48995066E+000     -5.48989189E+000  <-- R
N      1         1.15732344E-004      1.10453835E-004     -1.10452023E-004  <-- V
N      2        -1.15732344E-004     -1.10453835E-004      1.10452023E-004  <-- V
N      1        -1.83347496E-004      1.53896599E-003     -1.53886170E-003  <-- F
N      2         1.83347496E-004     -1.53896599E-003      1.53886170E-003  <-- F
```

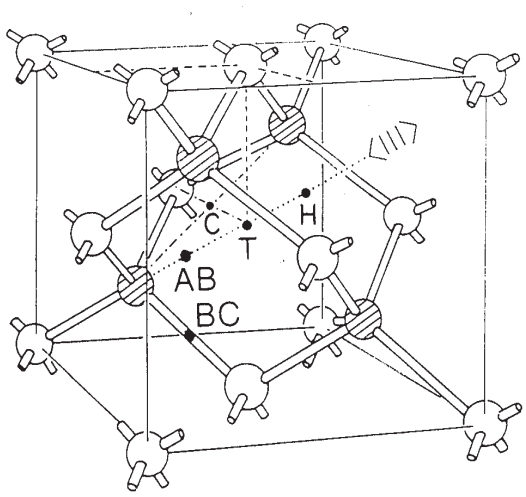

Visualisation

- PIMD produces usual CASTEP output PLUS
- `<seedname>_pimdXXX.md` file for each bead (1<= XXX <= P)
- These files are identical to normal `<seedname>.md` file but get 1 for all particles at same value of imaginary time.
- Can then use conventional CASTEP MD tools to analyse such as MDTEP
- Or use the `pi_merge` script to combine into a single file for visualisation (use md2xyz) ...

CASTEP PIMD
case studies

THE UNIVERSITY of York

Hydrogen in Silicon



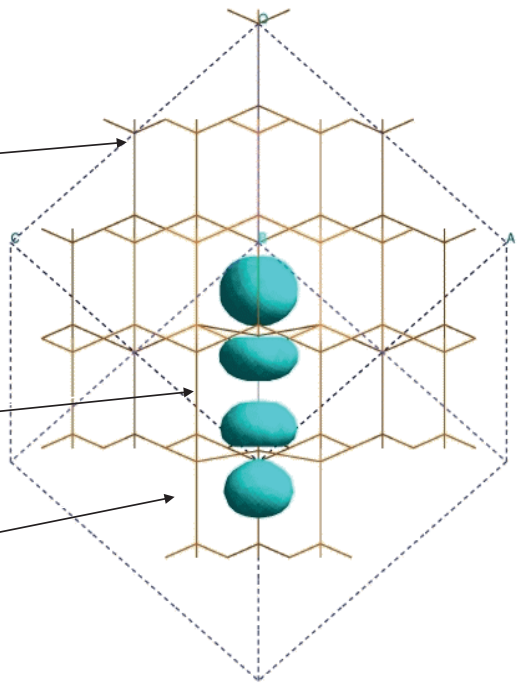
- Stable / Metastable sites
  - BC two-fold coordinated
  - T four-fold coordinated
- Possible saddlepoint sites
  - AB antibonding site
  - C half-way to T
  - H hexagonal (6-fold) site

Spin Density at BC site

Planes of silicon atoms seen "edge-on"

BC site

Spin density iso-surface due to single extra electron



T=0 Results

- Large lattice strain around BC site
- Small (inwards) relaxation around T
- Both sites stable with BC preferred to T
- Relative energy (BC–T) ~ -0.27eV

Bond	Length (Å)
Si-Si in bulk	2.351
Si-H in SiH4	1.480
Si-H at T	2.278 (-3%)
Si-H at BC	1.650 (+40%)

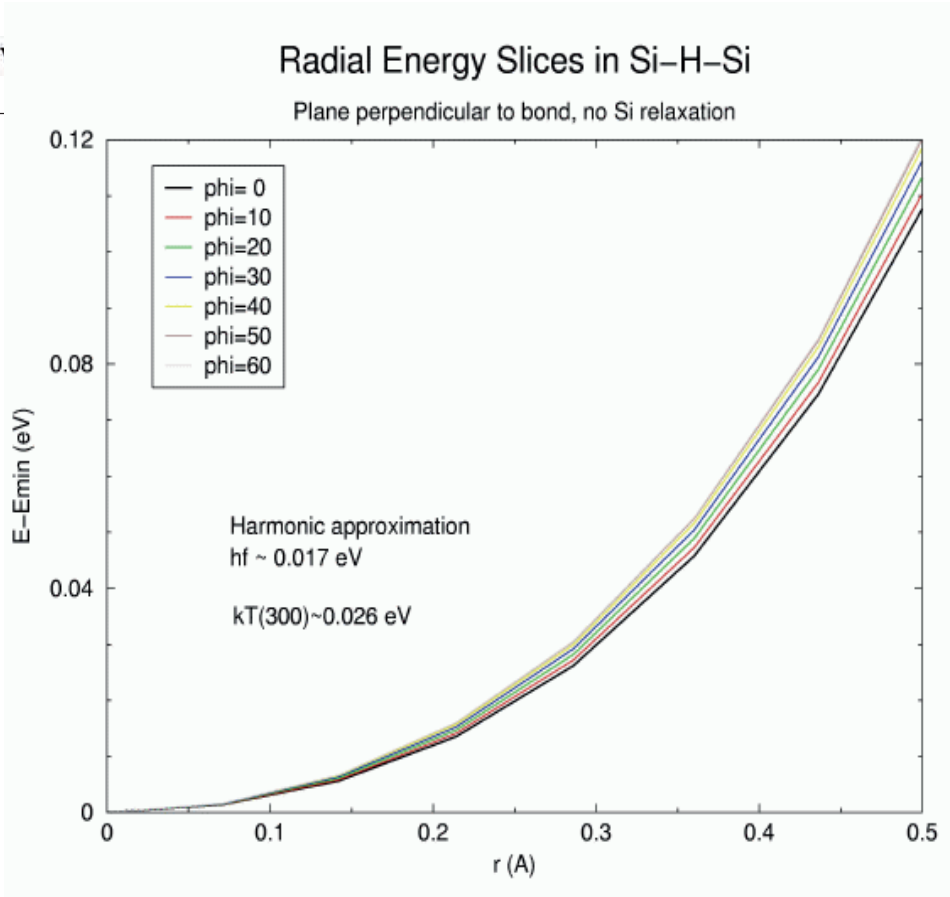
Site	Lattice Relaxation Energy (eV)
H at T	0.032
H at BC	1.662

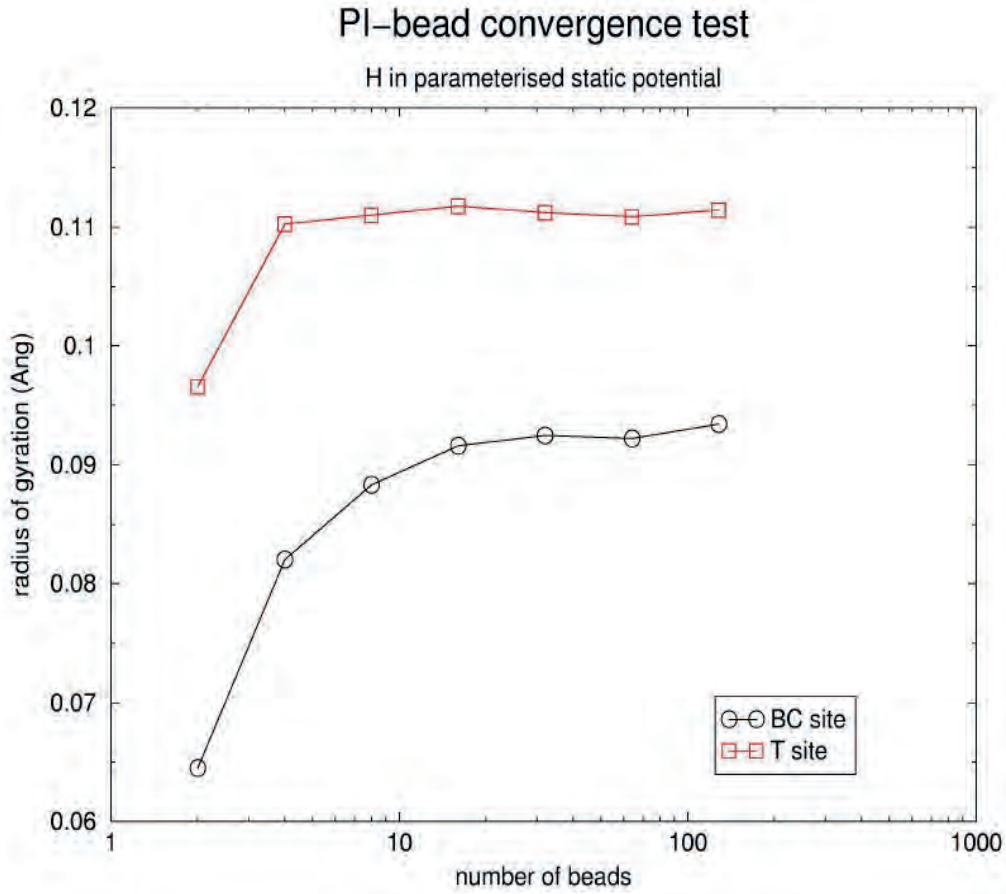
Site	Binding Energy (eV)
H at T	0.284
H at BC	0.554

Awkward Experimental Fact

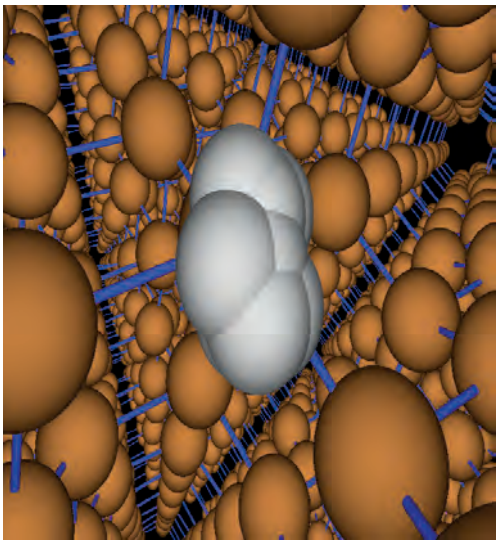
μSR → 1:2 population of BC:T sites …

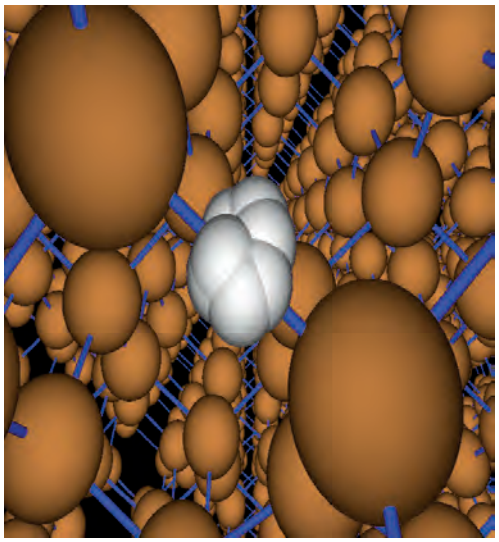
- BUT we see BC<T and there are 8 BC sites for every T site!
- Is it a thermal effect?
  - *Ab initio* MD suggests no significant energy changes
- Non-equilibrium effect?
  - need barrier heights → saddlepoints → yet to be tackled
- Is it a quantum effect?
  - Mass Mu ~ $^1/_9$ Mass H and ZPM ~ 1/sqrt(mass) …





Visualisations

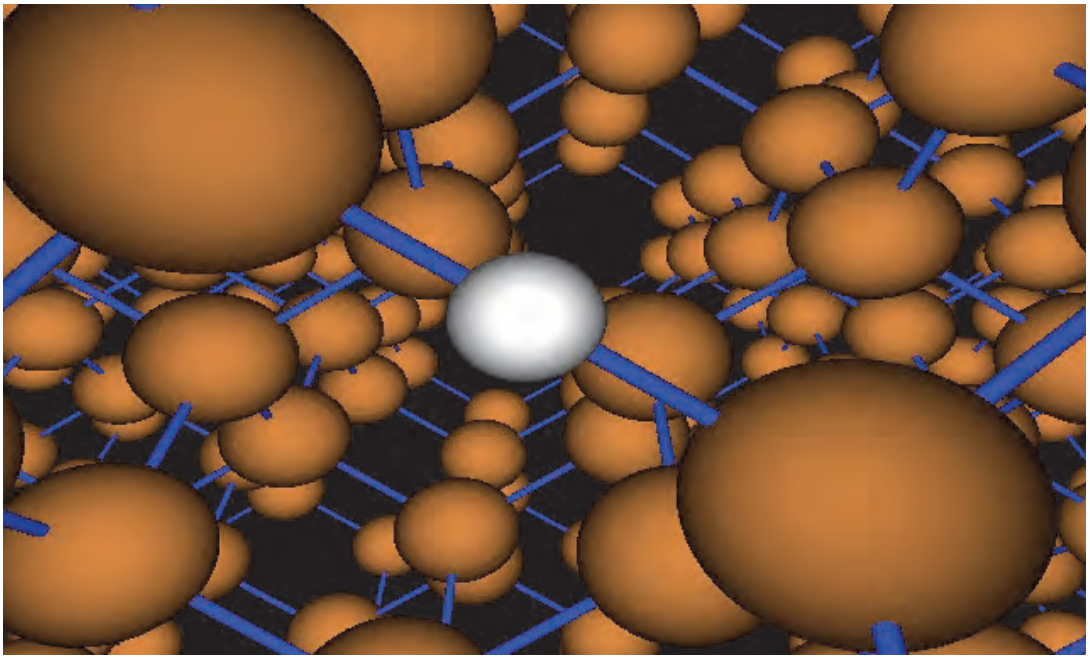




Superimposing all beads at same value of imaginary time
at a single instance of MD time (T=300 K, P=16).

THE UNIVERSITY of York

H-Defect at BC Site in Silicon



THE UNIVERSITY of York

Energy Results

- PI is indeed capturing the quantum effects
  - big difference in energies when turn PI on
  - now get relative energy (BC-T) ~ -0.08 eV
- Still have conventional view BC<T
- Adding ZPM increases energy at both sites
  - bigger effect at BC than T due to confinement
  - enhanced effect for Mu expected

Implications

- Hydrogen
  - BC is still confined <u>BUT</u> T may not be, *i.e.* no longer fixed but mobile/delocalized
  - need definitive saddlepoints
- Muonium
  - probable cross-over in ordering of sites
  - probably BC confined but T highly mobile
  - no longer a good analogue for Hydrogen?

H_2O on metal surfaces

Water-hydroxyl layers

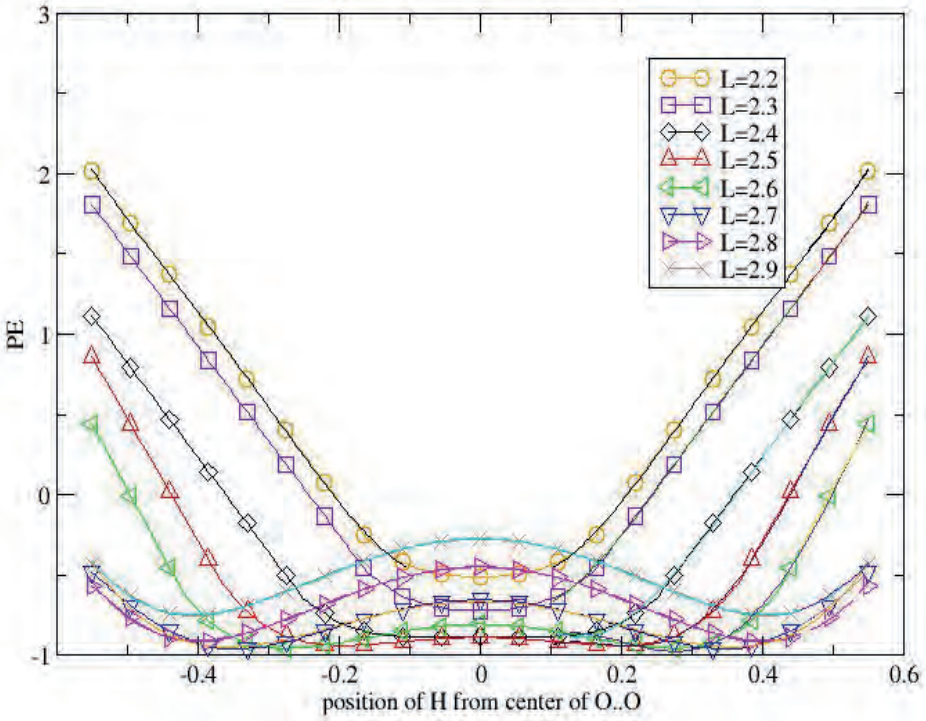
- In many systems, the initial wetting layer is not pure water, but a water-hydroxyl mix
  - Bond lengths/angles unusual due to “pinning” with hydrogen-bonds formed to surface atoms
- Transition metal surfaces have been well-characterised
  - Pt(111) has large lattice constant and so inter-molecule distance ~ 2.83 Å
  - Ni(111) has much smaller distance ~2.50 Å

“*Water-Hydroxyl Overlayers on Metal Surfaces*”
Phys. Rev. Lett. **104**, 066102 (2010)

Water problem

- In bulk ice have typical O-O distance ~ 2.8Å
- At high pressures (>70 GPa) ice has typical O-O distance of ~2.3 Å
  - No longer a molecular crystal
  - Have delocalised protons between O nuclei
- Low T (160 K) measurements of hydrogen diffusion on metal surfaces suggests that quantum tunnelling important
  - Hence need full QM treatment for hydrogen!
  - Short cut to converge number of beads …

Static H_2O/OH on Ni PES



THE UNIVERSITY of York

Nickel PIMD Movie

THE UNIVERSITY of York

Platinum PIMD Movie

THE UNIVERSITY of York

Hydrogen in Minerals

Hydrogen is a common component of many minerals but position very hard to locate by traditional techniques

Brucite - $Mg(OH)_2$ – is a simple mineral but location of H unclear

PIMD shows why …

Other examples

- Recent paper on high-pressure hydrogen phase diagram
  - Correcting previous paper on structure search to include ZPE and finite T
- Recent paper on high-pressure melting of hydrogen – two-phase coexistence with PIMD
- Paper under review – diffusion of H on Ru – looking at quantum vs classical diffusion vs T
- Current project – diffusion of H on Ni and isotope effects

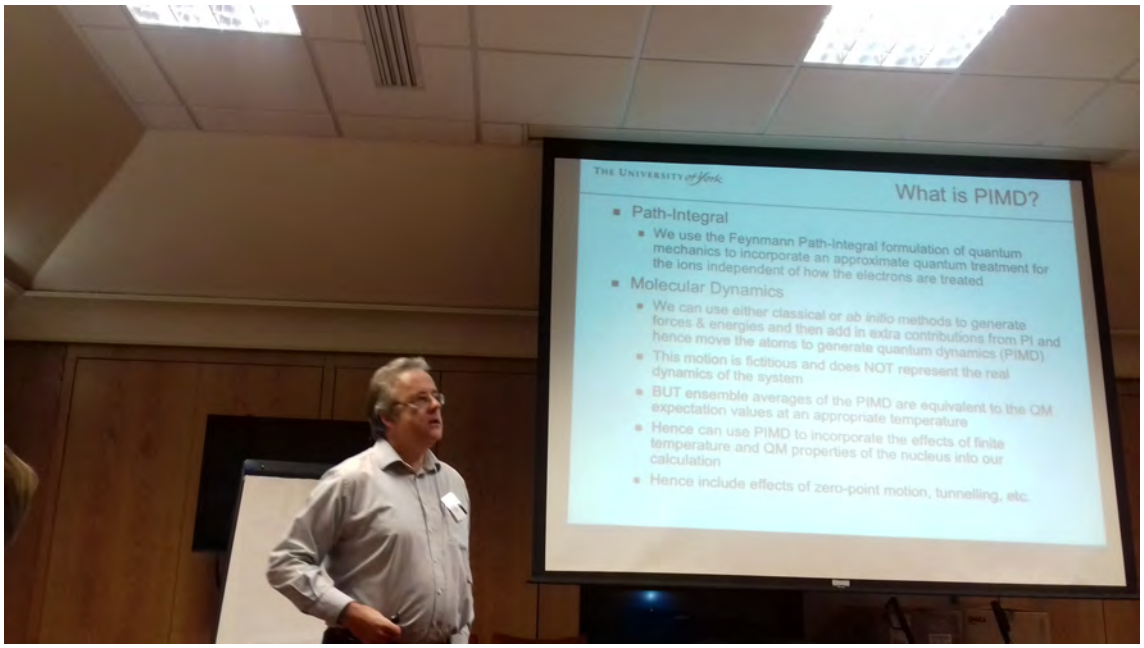
Summary

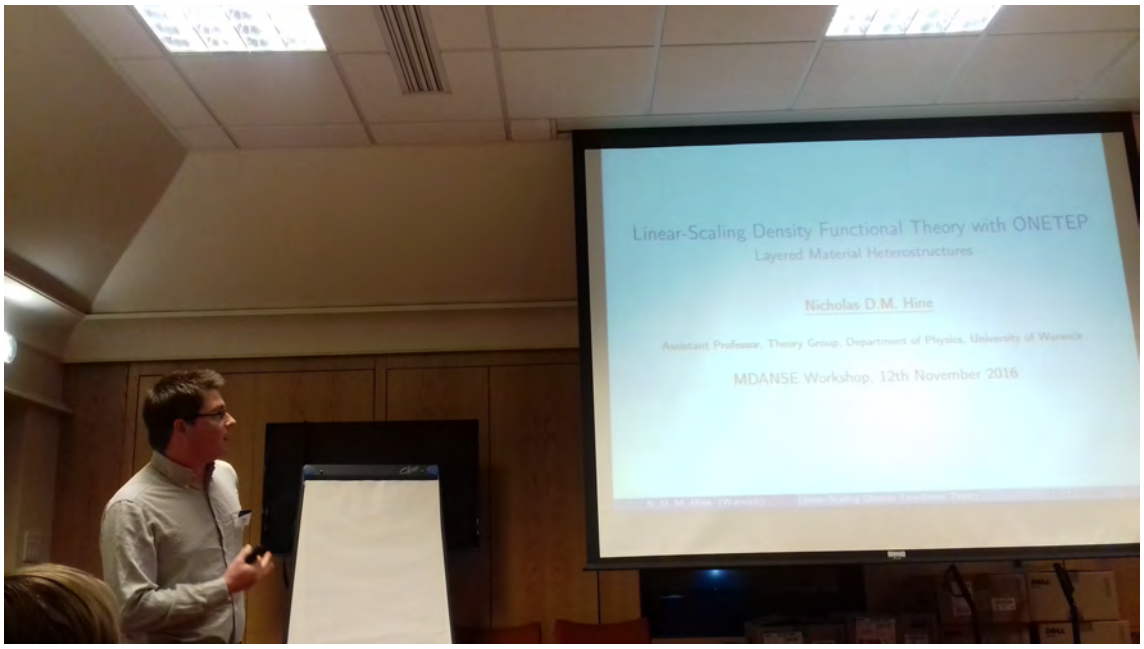
Summary

- PIMD as a way of going beyond BOMD
  - Quantum treatment of ions but expensive!
- Usual MD caveats
  - Beware equilibration, not all configurations are equal, consider sampling and correlation, etc.
- Apply basic physics to the results
  - conservation laws, equipartition, etc
- Additional concern
  - need to converge w.r.t. number of beads
- BEWARE:
  - the dynamics ARE FICTITIOUS and only the ensemble average is meaningful …
  - some recent theoretical developments (e.g. centroid PIMD, ring-polymer PIMD) can do this – PIMD still being actively developed…

References

- R.P. Feynman, Rev. Mod. Phys **20**(2) 367 (1948)
  - The original paper
- "*Quantum Mechanics and Path Integrals*"
  - R.P. Feynman and A.R. Hibbs (McGraw-Hill) (1965). Definitive Text.
- "*Computer Simulation of Liquids*"
  - M.P Allen & D.J. Tildesley (1987). Chapter 10 old but useful.
- M.J. Gillan, Phil. Mag. A **58**(1) 257 (1988)
  - semi-classical but nice description of PIMD of H in metals
- M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, J. Chem. Phys. **104**, 5579 (1996)
  - Efficiency improvements such as staging modes
- **www.castep.org** web site
  - Useful MD and geometry optimisation tutorials, plus FAQs, on-line keyword listing, MDTEP download, etc.


THE UNIVERSITY of York
What is PIMD?
Path-Integral
We use the Feynmann Path-Integral formulation of quantum mechanics to incorporate an approximate quantum treatment for the ions independent of how the electrons are treated
Molecular Dynamics
We can use either classical or ab initio methods to generate forces & energies and then add in extra contributions from PI and hence move the atoms to generate quantum dynamics (PIMD)
This motion is fictitious and does NOT represent the real dynamics of the system
BUT ensemble averages of the PIMD are equivalent to the QM expectation values at an appropriate temperature
Hence can use PIMD to incorporate the effects of finite temperature and QM properties of the nucleus into our calculation
Hence include effects of zero-point motion, tunnelling, etc.


Linear-Scaling Density Functional Theory with ONETEP
Layered Material Heterostructures
Nicholas D.M. Hine
Assistant Professor, Theory Group, Department of Physics, University of Warwick
MDANSE Workshop, 12th November 2016

Linear-Scaling Density Functional Theory with ONETEP

Layered Material Heterostructures

Nicholas D.M. Hine

Assistant Professor, Theory Group, Department of Physics, University of Warwick

MDANSE Workshop, 12th November 2016

Heterostructures of 2D Materials

Layered material in devices: many questions

Commensurate lattice?
Transfer vs Epitaxial Growth?
Effects of encapsulation on bandstructure?
Transport through full system, not just layer?
Band alignment, quality of contacts?
Stability in air?

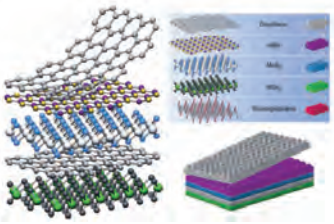
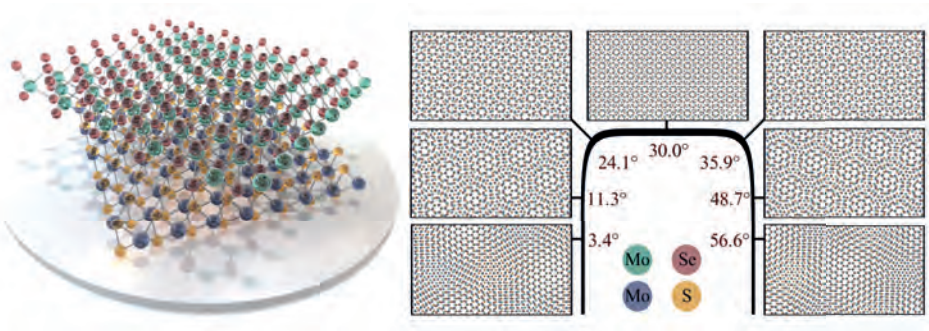


Image: A. K. Geim, I. V. Grigorieva, Nature 499, 419 (2013)

Misaligned / Incommensurate interfaces

Example: $MoS_2/MoSe_2$ heterostructure – 4-5% lattice mismatch produces incommensurate interface. Interlayer interactions may depend on lattice mismatch angle:



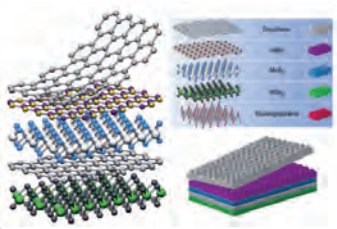
(images by G. Constantinescu)

How can we model these structures/devices and predict their properties?

Towards *Ab Initio* Device Modelling

Requirements for Underlying Method:

High-accuracy (DFT or equivalent)
Low-order scaling with system size
Cheap treatment of vacuum
van der Waals interactions
Spin-Orbit Coupling for heavier elements



Simulation Capabilities

Structure, Energetics
Theoretical Spectroscopy techniques
(phonons, optical, EELS, ARPES)
Electron Transport
Binding / Catalysis at surfaces

Traditional Density Functional Theory

Kohn-Sham DFT: functional of the electron density $\rho(\mathbf{r})$:

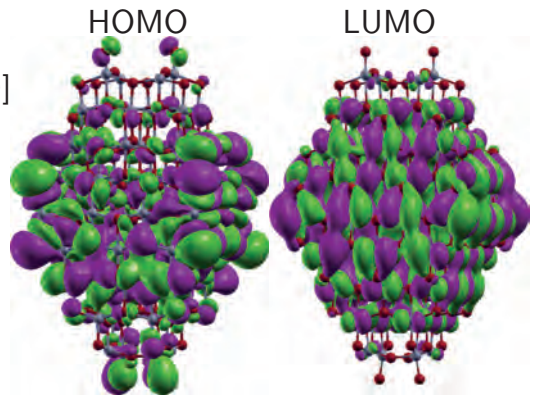
$$E_{KS}[\rho(\mathbf{r})] = T_s[\rho] + E_H[\rho] + E_{xc}[\rho] + E_{ext}[\rho]$$

Complexity of QM confined to $E_{xc}[\rho]$: reasonably simple yet powerful approximations

Minimise E_{KS} w.r.t. $\rho(\mathbf{r})$ with $\rho(\mathbf{r}) = \sum_n |\psi_n(\mathbf{r})|^2$, for orthogonal single-particle states $\psi_n(\mathbf{r})$

$$\hat{H}_{KS}[\rho(\mathbf{r})]\psi_n(\mathbf{r}) = \varepsilon_n \psi_n(\mathbf{r})$$

seeks eigenstates $\psi_n(\mathbf{r})$, eigenvalues ε_n.



Each orbital $\psi_n(\mathbf{r})$ is delocalised over whole system

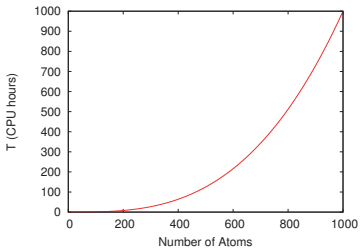
Traditional Density Functional Theory

Scalability Problem:

- For N atoms, eigenstate-based approaches are inherently $O(N^3)$:
  - Num. eigenstates $\propto O(N)$
  - Num. basis fns/eigenstate $\propto O(N)$
  - Each orthogonal to $\propto O(N)$ others

Becomes computational unfeasible much beyond around 400-800 atoms! Many systems (whole nanocrystals, grain boundaries, dilute defects, proteins, etc) need much more than this: 1000+ to enter realistic regime

Plane-wave codes like CASTEP, VASP, QE implement this in efficient, accurate and fairly scalable form.



Still $O(N^3)$ scaling of computational effort!

- Can we avoid eigenstates? Use density matrix, then exploit locality of QM in insulators

The Density Matrix

Single-electron density matrix contains all info required for evaluating operators:

$$\rho(\mathbf{r},\mathbf{r}') = \sum_n \psi_n(\mathbf{r}) f_n \psi_n(\mathbf{r}')$$

We could re-express our eigenstates in terms of local orbitals:

$$\psi_n(\mathbf{r}) = \sum_\alpha \phi_\alpha(\mathbf{r}) M^\alpha{}_n$$

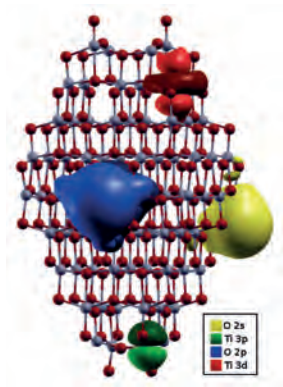
This transforms the D.M. as:

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta(\mathbf{r}')$$

Density kernel $K^{\alpha\beta} = \sum_n M^\alpha{}_n f_n (M^\dagger)_n{}^\beta$ is a generalisation of the occupation number.

NB: Density $\rho(\mathbf{r}) = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta(\mathbf{r})$

Can we exploit nearsightedness of QM in insulators? $\rho(\mathbf{r},\mathbf{r}') \to 0$ as $|\mathbf{r}-\mathbf{r}| \to \infty$



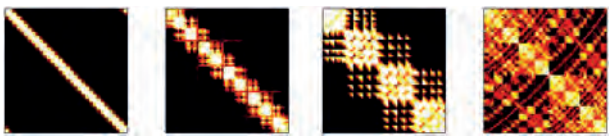
Therefore we can set $K^{\alpha\beta} = 0$ for $|\mathbf{R}_\alpha - \mathbf{R}_\beta| > R_K$

Sparse Matrices

- With local orbitals, we can evaluate Hamiltonian for any pair α, β:

$$H_{\alpha\beta} = \langle \phi_\alpha | \hat{H} | \phi_\beta \rangle$$

  Matrix is nonzero only if $\phi_\alpha(\mathbf{r})$ and $\phi_\beta(\mathbf{r})$ are overlapping.
- Same is true for the overlap matrix $S_{\alpha\beta} = \langle \phi_\alpha | \phi_\beta \rangle$. They are *sparse matrices*
- Density matrix *nearsighted* for insulators, so $K^{\alpha\beta}$ is *sparse*.
- Matrix products such as $\sum_\beta K^{\alpha\beta} S_{\beta\gamma}$ are still sparse and can be calculated in $O(N)$



Left→Right: CNT, DNA, GaAs Nanorod, Bulk-Si ($\sim$ 4000 atom systems)

$O(N)$ approaches to DFT

- Bandstructure energy in KS approach given by

$$\begin{aligned} E_{bs} &= \sum_n f_n \varepsilon_n = \sum_n f_n \langle \psi_n | \hat{H} | \psi_n \rangle \\ &= \sum_n f_n \sum_{\alpha\beta} (M^\dagger)_n{}^\beta \langle \phi_\beta | \hat{H} | \phi_\alpha \rangle M^\alpha{}_n \\ &= \sum_{\alpha\beta} K^{\alpha\beta} H_{\beta\alpha} \end{aligned}$$

  This is the trace (sum of diagonal elements) of the matrix **K.H**

- We need to check if the density matrix contains the right number of electrons:

$$\begin{aligned} N_e &= \int \rho(\mathbf{r}) \mathrm{d}\mathbf{r} \\ &= \int \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta(\mathbf{r}) \mathrm{d}\mathbf{r} \\ &= \sum_{\alpha\beta} K^{\alpha\beta} S_{\beta\alpha} \end{aligned}$$

  This is the trace of the matrix **K.S**. We have $\mathrm{Tr}(\mathbf{K.S}) = N_e$ as a *constraint*

$O(N)$ approaches to DFT

- We also need to maintain crucial property of *idempotency*, which results from mutual orthogonality of eigenstates (as long as occupancies f_m are 0 or 1):

$$\begin{aligned} \rho^2(\mathbf{r},\mathbf{r}') &= \int \rho(\mathbf{r},\mathbf{r}'') \rho(\mathbf{r}'',\mathbf{r}') \mathrm{d}\mathbf{r}'' \\ &= \sum_{nm} \psi_n(\mathbf{r}) f_n \int \psi_n(\mathbf{r}'') \psi_m(\mathbf{r}'') \mathrm{d}\mathbf{r}'' f_m \psi_m(\mathbf{r}') \\ &= \sum_{nm} \psi_n(\mathbf{r}) f_n \delta_{nm} f_m \psi_m(\mathbf{r}') = \rho(\mathbf{r},\mathbf{r}') \end{aligned}$$

- Expressing this in terms of local orbitals and sparse matrices gives us:

$$\begin{aligned} \rho(\mathbf{r},\mathbf{r}') &= \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta(\mathbf{r}) \\ \rho^2(\mathbf{r},\mathbf{r}') &= \phi_\alpha(\mathbf{r}) K^{\alpha\gamma} \int \phi_\gamma(\mathbf{r}'') \phi_\delta(\mathbf{r}'') \mathrm{d}\mathbf{r}'' K^{\delta\beta} \phi_\beta(\mathbf{r}') \\ K^{\alpha\beta} &= K^{\alpha\gamma} S_{\gamma\delta} K^{\delta\beta} \qquad \text{constraint to enforce idempotency} \end{aligned}$$

$O(N)$ approaches: Ingredients

- Minimise energy, enforcing normalisation and idempotency:

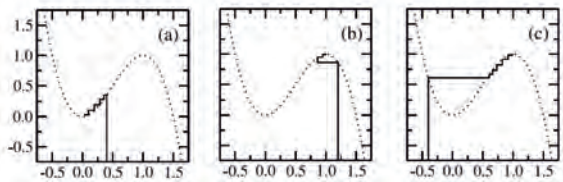
$$\text{Energy:} \quad E_T = K^{\alpha\beta} H_{\beta\alpha} - E_{\text{dc}}$$

$$\text{Normalisation:} \quad N_e = K^{\alpha\beta} S_{\beta\alpha}$$

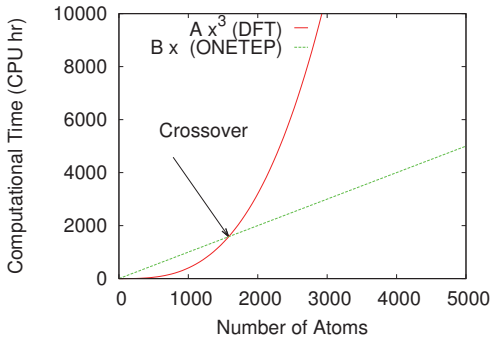
$$\text{Idempotency:} \quad K^{\alpha\beta} = K^{\alpha\gamma} S_{\gamma\delta} K^{\delta\beta}$$

- Optimisation algorithms (e.g. Li Nunes Vanderbilt) incorporating McWeeny purification transformation:

$$f_{new} = 3f_{old}^2 - 2f_{old}^3$$



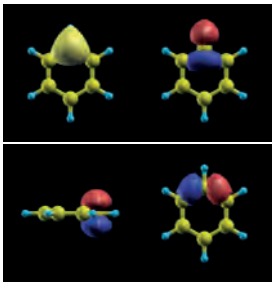
- Efficient, flexible and accurate local orbital basis
- Need to evaluate density, Hamiltonian matrix and overlap matrix in $O(N)$ computational effort
- Algorithms to optimise density matrix using sparse matrix algebra



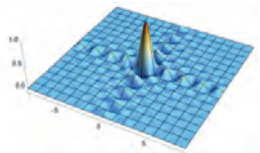
$O(N)$ algorithms have higher prefactor than $O(N^3)$ so only favourable after "crossover"

ONETEP: Linear Scaling DFT

NGWFs

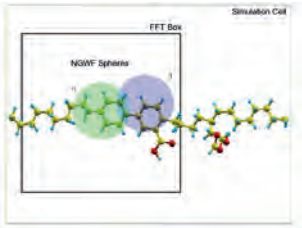


Minimal basis of local orbitals optimised in-situ, systematic wrt real & recip-space cutoffs.
Expressed in psinc functions:



FFT Box Approximation

Local $\phi_\alpha \Rightarrow$moving FFT box



Many uses throughout code:
K.E. matrix elements

NL-PSP Projectors

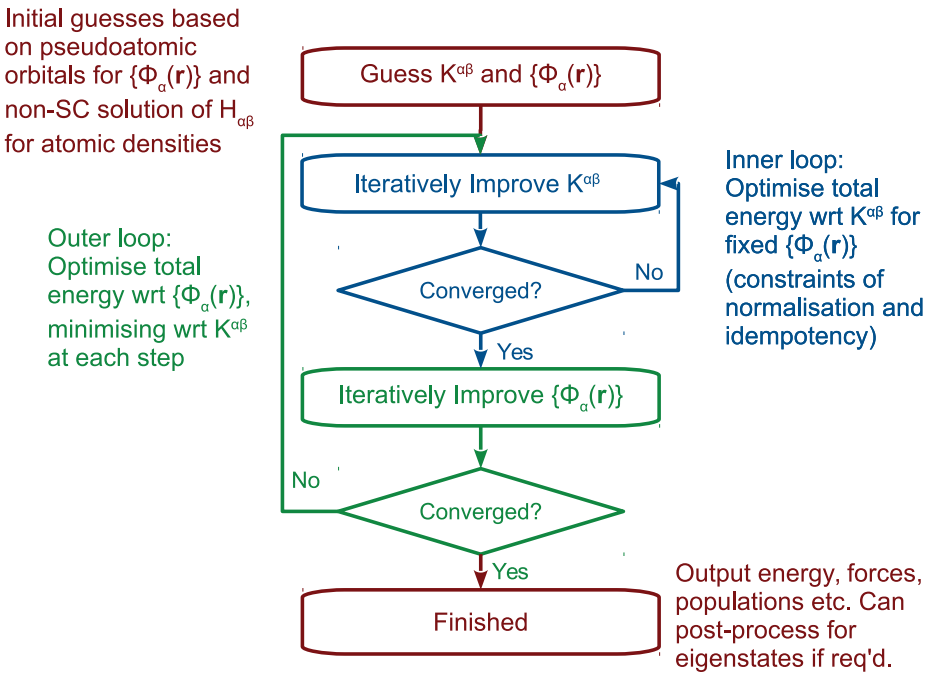
Interpolation of products $\phi_\alpha(\mathbf{r})\phi_\beta(\mathbf{r})$ to fine grid $n(\mathbf{r})$

Evaluation of integrals $\langle \phi_\alpha | V_{\text{eff}}(\mathbf{r}) | \phi_\beta \rangle$

$\Rightarrow$ **Strictly $O(N)$ evaluation of $H_{\alpha\beta}$ and $n(\mathbf{r})$**

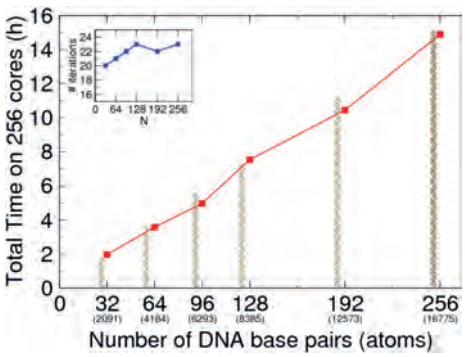
www.onetep.org; CK Skylaris, PD Haynes, AA Mostofi and MC Payne, J. Chem. Phys. 122, 084119 (2005)
NDM Hine, PD Haynes, AA Mostofi, C-K Skylaris, MC Payne, Comput. Phys. Commun. 180, 1041 (2009)

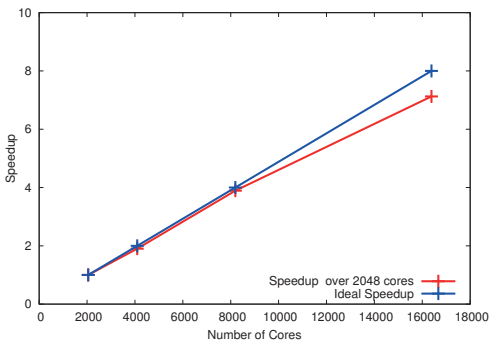
ONETEP: Linear Scaling DFT



ONETEP: Linear-Scaling DFT

Linear-Scaling computational effort: scales to tens of thousands of atoms





MPI/OpenMP Hybrid Parallelism - scales to tens of thousands of parallel cores.
Obtainable through (cheap) academic license, or by (free) collaborator agreement for specific collaborative projects.

van der Waals Density Functionals

- Semi-local density functionals (LDA, GGA, etc) lack van der Waals
  $\Rightarrow$ poor energetics and geometries for layered systems, organics, nanostructures.
- Fully non-local density functionals, eg Dion *et al* are challenging to evaluate, but Soler's method allows calculation in $O(N \log N)$:

$$E_c^{\mathrm{nl}}[\rho(\mathbf{r})] = \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}'\rho(\mathbf{r})\phi(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}')$$
$$= \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}'\rho(\mathbf{r})\rho(\mathbf{r}')\phi(q,q',|\mathbf{r}-\mathbf{r}'|) \qquad [q = q(\rho(\mathbf{r}),|\nabla\rho(\mathbf{r})|)]$$

$\phi(q,q',r)$ is smooth as fn of q so can be represented using ~ 20 interpolating polynomials: $\phi(q,q',r) = \sum_{\alpha\beta}\phi_{\alpha\beta}(r)p_\alpha(q)p_b(q')$. Defining $\theta_\alpha(\mathbf{r}) = \rho(\mathbf{r})p_\alpha(q(\mathbf{r}))$ gives

$$E_c^{\mathrm{nl}}[\rho(\mathbf{r})] = \frac{1}{2}\sum_{\alpha\beta}\int\int d\mathbf{r}d\mathbf{r}'\,\theta_\alpha(\mathbf{r})\phi_{\alpha\beta}(|\mathbf{r}-\mathbf{r}'|)\theta_\beta(\mathbf{r}')$$
$$= \frac{1}{2}\int\int d\mathbf{k}\,\theta_\alpha^*(\mathbf{k})\phi_{\alpha\beta}(k)\theta_\beta(\mathbf{k})$$

So can evaluate $E_c^{\mathrm{nl}}[\rho(\mathbf{r})]$ with just a handful of whole-cell FFTs.

Transition Metal Dichalcogenides: Geometry

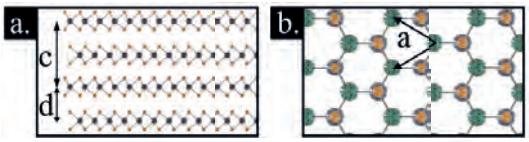
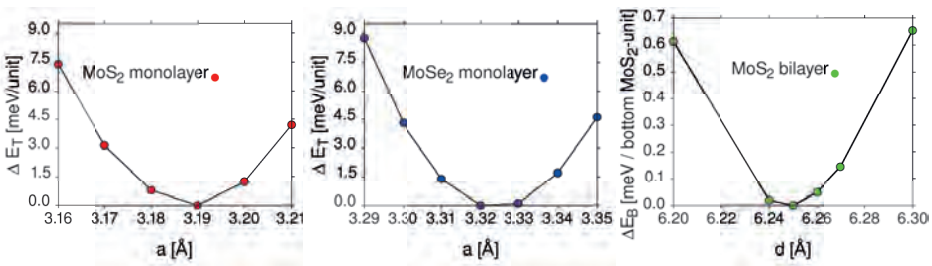


TABLE I. Bulk MoS_2 lattice optimisation

	optB88	optPBE	vdw-DF	vdw-DF2	Exp.
a[Å]	3.19	3.20	3.24	3.29	3.16
c[Å]	12.50	12.74	13.14	12.91	12.30

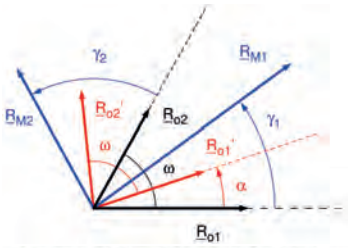
- Bulk calculations: plane-wave DFT, Quantum ESPRESSO
- OPTPBE/OPTB88 vdW-DF functionals (Klimes *et al*, 2011): interlayer distance in good comparison to experiment for bulk MoS_2
- OPTB88 used for the rest of these results

- Monolayer and bilayer primitive cell geometries can then be calculated:



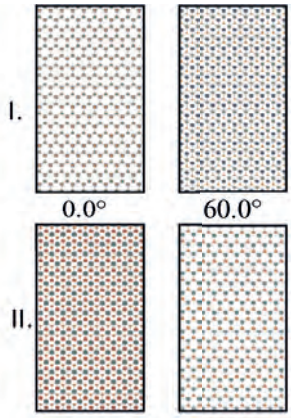
Transition Metal Dichalcogenides: LS-DFT Calculations

Magic angles ⇒ very small cells with commensurate lattice.
[Occurs only for either a homostructural bilayer or when strain forces equal lattice constants.]



Binding energy (meV/bottom MoS_2-unit) for twisted MoS_2 bilayers

Angle [°]	case I (0.0°)	case II (0.0°)
0.0	-220.1	-207.3
60.0	-219.0	-65.9



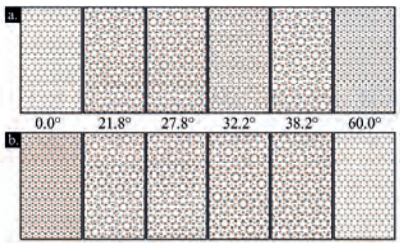
For other alignments, try to find supercell for one layer for which low-strain combinations of lattice vectors of other material are equal.

Transition Metal Dichalcogenides: LS-DFT Calculations

Low-strain structures ⇒1000-4000 atom supercells

Binding energy (meV/bottom MoS_2-unit) for twisted MoS_2 bilayers

Angle [°]	case I (0.0°)	case II (0.0°)
0.0	-220.1	-207.3
21.8	-172.2	-172.3
27.8	-172.4	-172.4
32.2	-172.4	-172.4
38.2	-172.2	-172.3
60.0	-219.0	-65.9



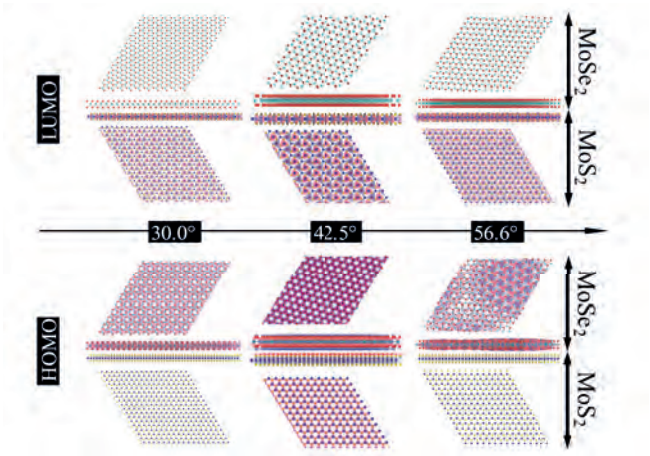
Indicates that apart from at magic angles, there is minimal variation in the binding energy with either angle or translation

⇒No need to treat in-plane displacements for non-magic angles

(sampled implicitly by varying alignment over the large supercells)

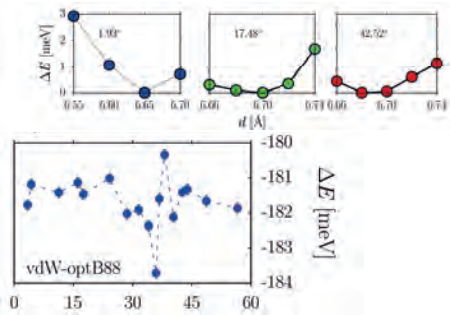
Transition Metal Dichalcogenides: LS-DFT Calculations

Keep strain under 1%, $\Rightarrow N \sim 1000 - 2000$: generate ensemble of geometries for $MoS_2/MoSe_2$ interface

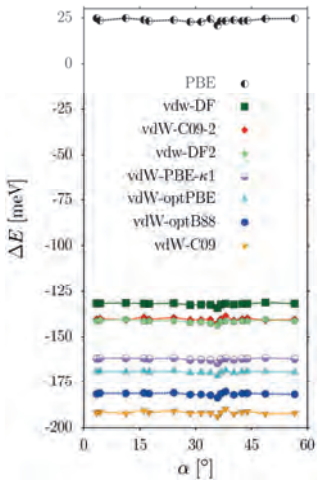


- Spatial separation of overall VBM and CBM of system on different materials
- Confirmed experimentally: F. Ceballos *et al*, ACS Nano 8, 12717 (2014).
- Moiré patterns affect localisation of orbitals — suggests impact on bandstructure

TMDC Heterostructures: Binding Energies



- Unlike in MoS_2 bilayers, minimal variation of distance with alignment angle:
- Crucially, binding energy is not sensitive to angular alignment of layers
- Neither perturbation theory nor momentum matching condition $N = \sum_{\mathbf{G_1},\mathbf{G_2}} \delta_{\mathbf{G_1},\mathbf{G_2}}$ can predict these results



- Different functionals $\Rightarrow$ different binding energies (expt unknown)
- All vdW-DFs produce similar interlayer distances, minimal variation with angle (< 5meV)

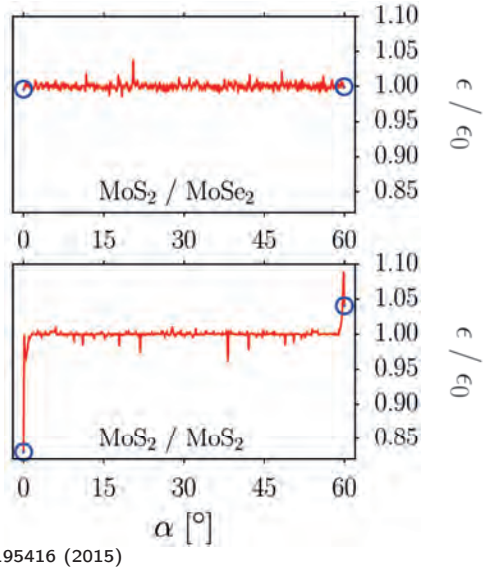
G. C. Constantinescu, N. D. M. Hine, Phys Rev B 91, 195416 (2015)

TMDC Heterostructures: Binding Energies

Universal shape of curves suggest remaining variations are all electrostatic in origin: can estimate possibility of binding with local density-potential perturbation model:

$$\varepsilon = \int \rho_1(\mathbf{r}) V_2(\mathbf{r}) d\mathbf{r} + \int \rho_2(\mathbf{r}) V_1(\mathbf{r}) d\mathbf{r}$$

Indicates that heterostructure lattice mismatch eliminates high-coupling angles due to momentum-matching condition not being fulfilled for small enough values of **G**.
Range of variation of energy estimated at $\sim$ 6meV/f.u., in agreement with full LS-DFT calculations.



G. C. Constantinescu, N. D. M. Hine, Phys Rev B 91, 195416 (2015)

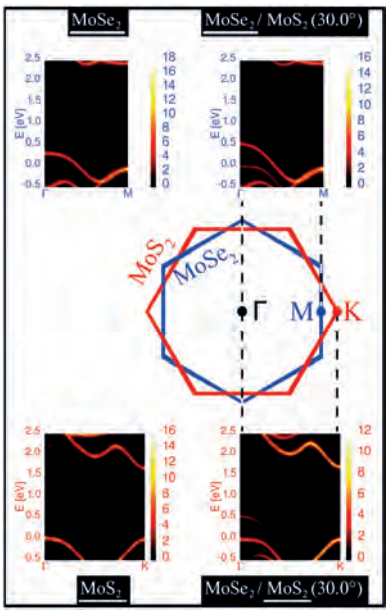
TMDC Heterostructures: Bandstructure Projection

- Despite decoupling of total energies, heterostructure construction affects individual bands & may strongly affect electronic properties
- Investigate by unfolding bands with a projector to the local orbitals of each layer:

$$A^{l}_{\mathbf{k}j,\mathbf{k}j} = \langle \psi_{\mathbf{k}j} | \left[\sum_{KJ} |\Psi_{KJ}\rangle A_{KJ,KJ} \langle \Psi_{KJ} | \right] \hat{P}_L | \psi_{\mathbf{k}j} \rangle$$

$$\hat{P}_L = \sum_{\alpha \in L} |\phi_\alpha\rangle\langle\phi^\alpha| \qquad L = \text{layer } 1,2$$

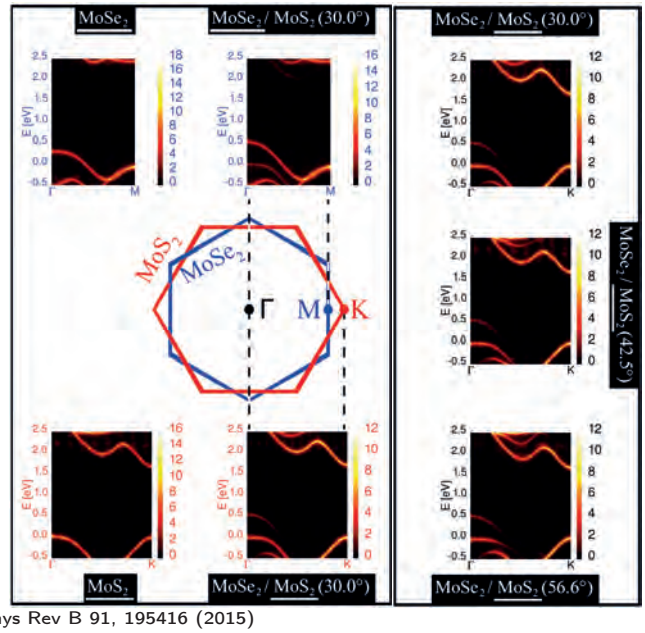
- Allows investigation of bands in each layer's primitive cell: evidence of hybridisation and intrusion of states of one layer into the other, leading to band repulsion.



G. C. Constantinescu, N. D. M. Hine, Phys Rev B 91, 195416 (2015)

TMDC Heterostructures: Bandstructure Projection

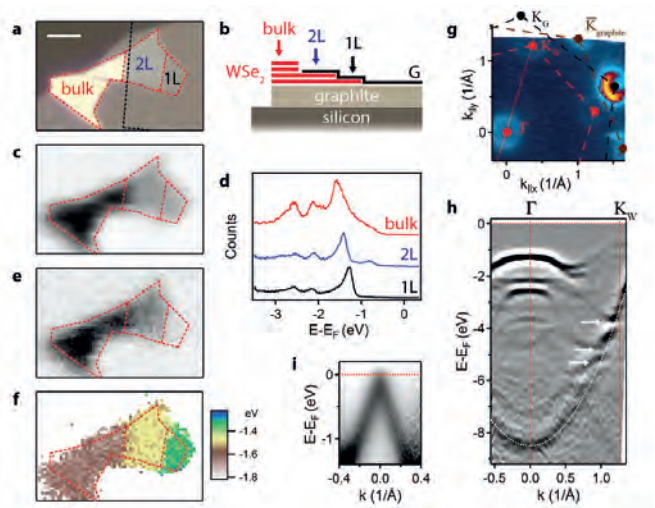
- Reduces spectral weight for $\Gamma \to K$ transitions, allowing for improved Photoluminescence.
- Indirect transitions redshifted compared to monolayers, as shown in MoS_2/WSe_2 bilayers
- Hole mass near Γ varies with angle (increases towards 60°) due to different portions of the BZ coming to momentum alignment.



G. C. Constantinescu, N. D. M. Hine, Phys Rev B 91, 195416 (2015)

Graphene-encapsulated WSe_2

μ-ARPES: Direct experimental probe of bandstructure

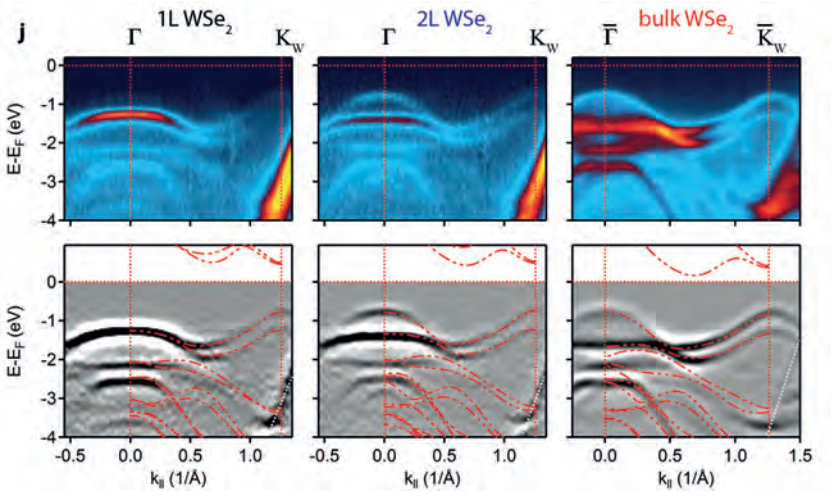


a. Optical image: Scale bar 5 μm;
b. Schematic cross-section
c. SPEM image (integrated within 3.5 eV of E_F).
d. Angle-integrated spectra from regions in a.
e. SPEM image in the energy window -1 eV.
f. Map of the energy of maximum emission, showing contrast between 1L, 2L and bulk regions.
g. Energy slice from full spectrum in 1L region showing BZ corners
h. Momentum slice $\Gamma - K_W$: Anticrossings between the graphene valence band and monolayer WSe2 bands

Figures by Neil Wilson (Warwick)

WSe_2 monolayer, bilayer and bulk

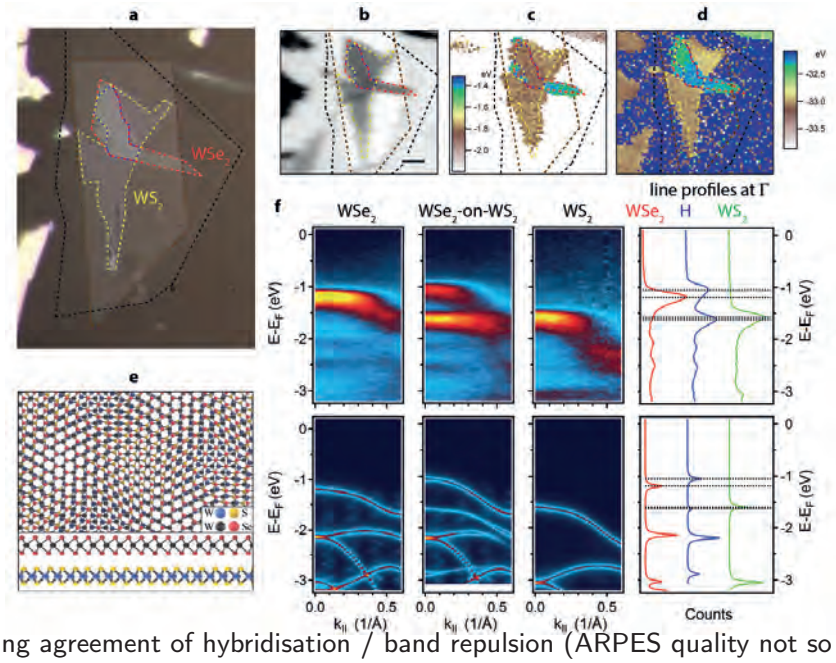
μ-ARPES agrees very well with DFT for states near VBM (1 parameter to align vertically)



Spin orbit splittings Δ_{so} particularly accurate

Rotated WSe_2 / WS_2 heterobilayer

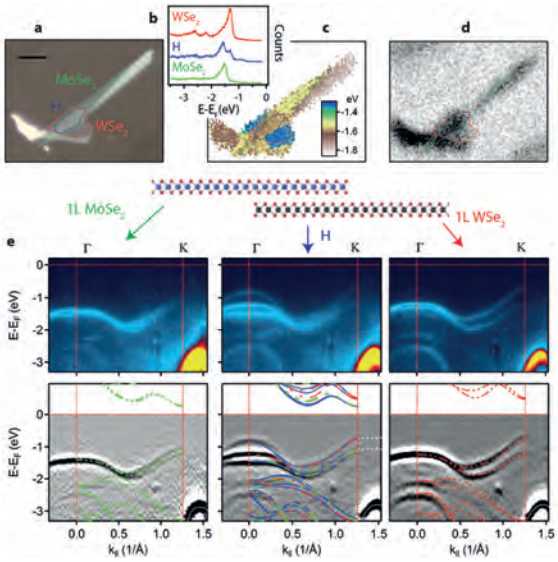
Superimpose two flakes to create WSe_2 / WS_2 interface (lattice mismatched)



Convincing agreement of hybridisation / band repulsion (ARPES quality not so good here)

Near 0° WSe_2 / $MoSe_2$ heterobilayer

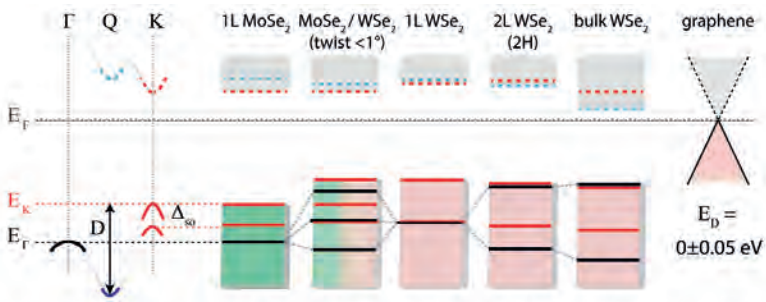
Flakes can be aligned at 0°/60° by aligning long edges. WSe_2 / $MoSe_2$ have very low lattice mismatch.



Heterostructure region clearly shows **three** bands at Γ! Possibly due to mixed commensurate/incommensurate regions.

Establishing Accurate Band Alignments

Measured band parameters (Solid lines: ARPES; dotted lines DFT).



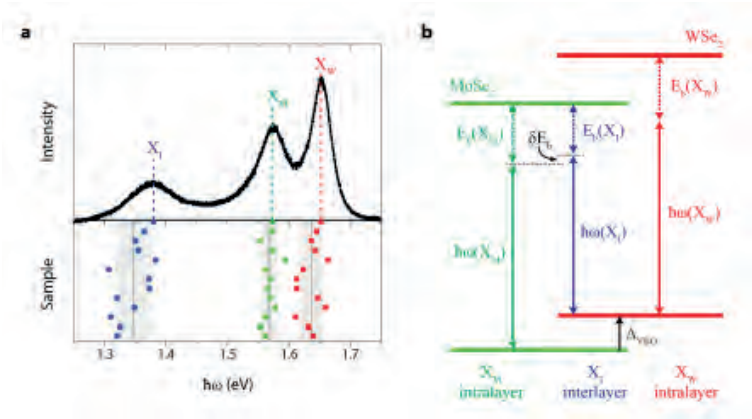
	1L $MoSe_2$	$MoSe_2$ / WSe_2 (twist <1°)	1L WSe_2	2L WSe_2 (2H)	bulk WSe_2	graphene
E_K (eV)	-0.99	-0.70	-0.70	-0.74	-0.79	±0.05
E_K-E_Γ (eV)	0.44	0.14	0.50	0.04	-0.04	±0.03
Δ_{so} (eV)	0.24	-	0.49	0.50	0.50	±0.03
D (eV)	0.90	1.25	1.24	1.20	1.24	±0.03
m^*_K / m_0	-	0.4	0.4	0.5	0.4	±0.1
m^*_Γ / m_0	4.2	1.2	2.8	1.9	0.9	±0.2

In both 2L WSe2 and heterobilayer MoSe2/WSe2, hybridization is almost undetectable at K (red bands) but much larger at (black bands).

N. R. Wilson, ..., G. C. Constantinescu, N. D. M. Hine, arXiv:1601.05865 (2016)

Direct Measurement of interlayer exciton

Measured photoluminescence for aligned MoSe2/WSe2 heterobilayer



Intralayer (X_M and X_W) and interlayer (X_I) excitons (2.33 eV excitation at 20 W). X_I is 220±20 meV below X_M.

N. R. Wilson, ..., G. C. Constantinescu, N. D. M. Hine, arXiv:1601.05865 (2016)

Tunnel FET

- Traditional MOSFET: voltage modulates thermionic emission over a barrier.
- Tunnel FET: voltage directly modulates quantum tunnelling.

TFETs not limited by the thermal Maxwell–Boltzmann tail of carriers, which limits subthreshold swing of MOSFETs to 60 mV/dec at room temperature

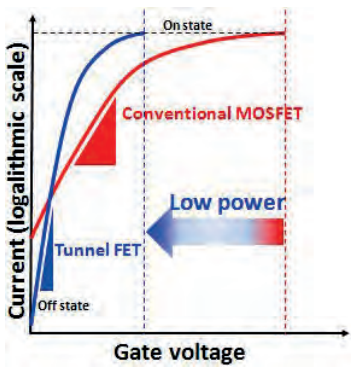
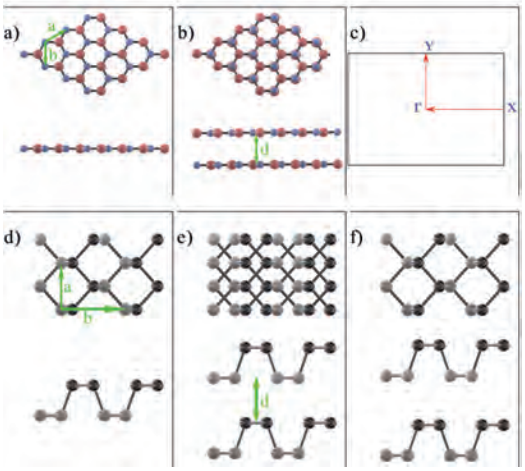


image source: phys.org

Black Phosphorous / h-BN



Hexagonal Boron Nitride:

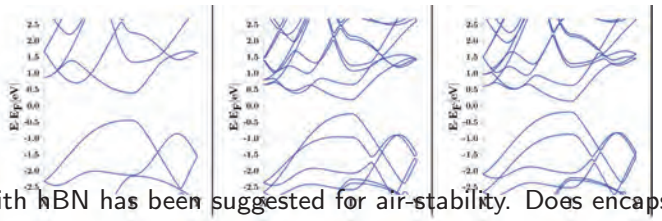
- Very wide bandgap, very stable
- VBM, CBM well away from those of black phosphorous
- Can be made very nearly defect-free
- Good electrical insulator

Black Phosphorous:

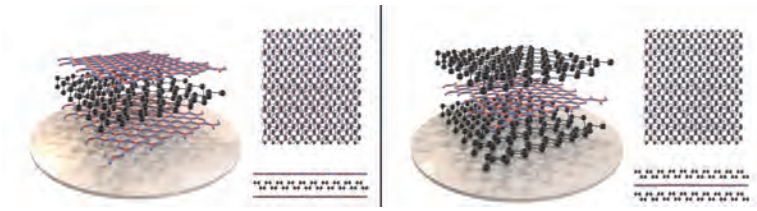
- Bandgap in useful range
- Not stable in air!

Black Phosphorous / h-BN

Multilayer (even bilayer) BP has greatly reduced bandgap compared to monolayer

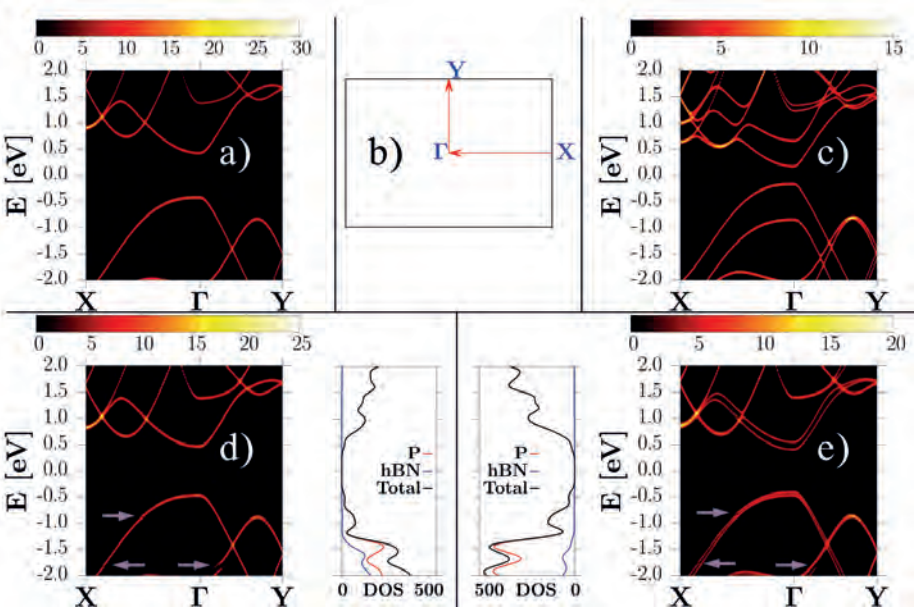


Encapsulation with hBN has been suggested for air-stability. Does encapsulation with h-BN affect electronic structure of BP? Is there interlayer interaction between monolayers of BP separated by h-BN?



Four model systems:
a) monolayer BP; b) hBN/BP/hBN; c) bilayer BP; d) BP/hBN/BP;

Spectral Function projections



G. C. Constantinescu, N. D. M. Hine, Nano Letters, 16, 2586 (2016)

ml-hBN/BP/hBN/BP/ml-hBN as a TFET

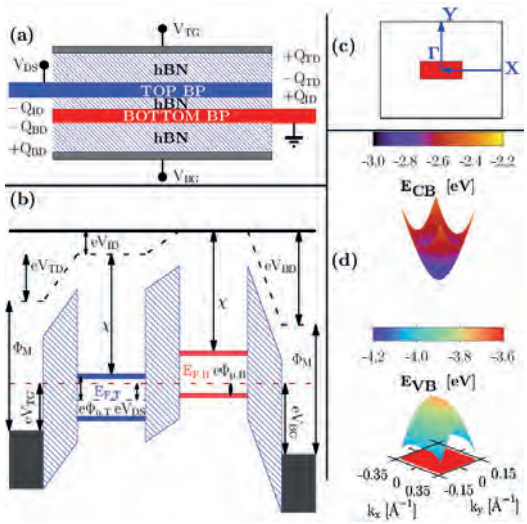
Tunnelling Current from T to B layers, summed over bands i,j, momenta $\mathbf{k}_T$, $\mathbf{k}_B$

$$I = \frac{4\pi e}{\hbar} \sum_{i,j} \sum_{\mathbf{k}_B,\mathbf{k}_T} |O_{ij}(\mathbf{k}_B,\mathbf{k}_T)|^2 \, \delta\left[E_{iB}(\mathbf{k}_B) - E_{jT}(\mathbf{k}_T)\right] \left(f_{iB} - f_{jT}\right)$$

Can be restricted to reciprocal-space region where there is only one peak in valence band and one valley in conduction band, i.e. red shaded area

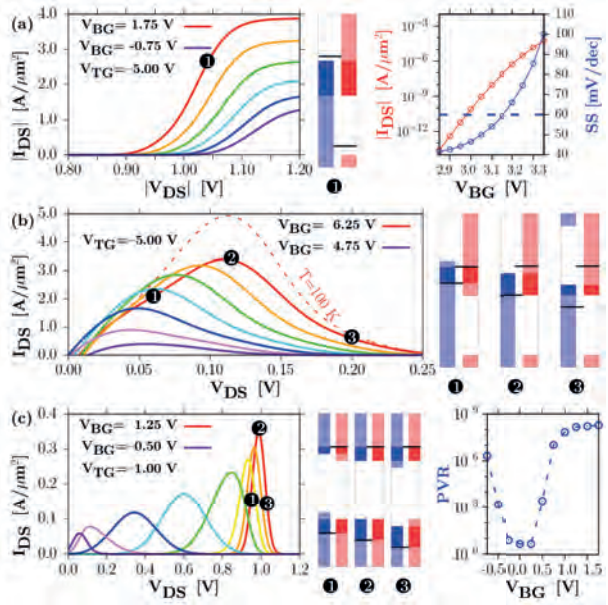
$$O_{ij}(\mathbf{k}_B,\mathbf{k}_T) = \langle \tilde{\psi}_{iB}(\mathbf{k}_B) \mid V \mid \tilde{\psi}_{jT}(\mathbf{k}_T) \rangle$$

Solve system of nonlinear equations for charges Q_{AB} in terms of e, h populations, DOS, voltages & Fermi levels.



G. C. Constantinescu, N. D. M. Hine, Nano Letters, 16, 2586 (2016)

ml-hBN/BP/hBN/BP/ml-hBN as a TFET



Three possible operating regimes:

a) Zener Diode in reverse bias
b) Broken gap arrangement:
Electrons in one layer tunnel to hole states in other.
c) Aligned gap arrangement:
Electrons ⇔ Electrons
Holes ⇔ Holes.

Regimes b), c) exhibit *Negative differential resistance* (NDR): for some part of I–V curve current decreases as the voltage increases. High Peak-to-Valley ratios achievable at relatively low voltages.

G. C. Constantinescu, N. D. M. Hine, Nano Letters, 16, 2586 (2016)

Recent functionality additions

- Linear-response TDDFT for excited states
- Implicit Solvent methods
- Constrained DFT for CT states
- Electron Energy Loss Spectroscopy
- FD and Linear-response phonons

Conclusions

- Linear Scaling DFT with ONETEP (www.onetep.org) allows plane-wave DFT simulations of layered material systems comprising thousands of atoms eg heterostructures
- Binding energies of lattice mismatched structures is nearly independent of alignment angle
- Strong angle-dependence of bandstructure effects: tunable effective mass and band offset
- Ab Initio device modelling possible for candidate TFET structures
- Advanced MD functionalities (eg asnharmonic infrared spectroscopy)

Acknowledgements

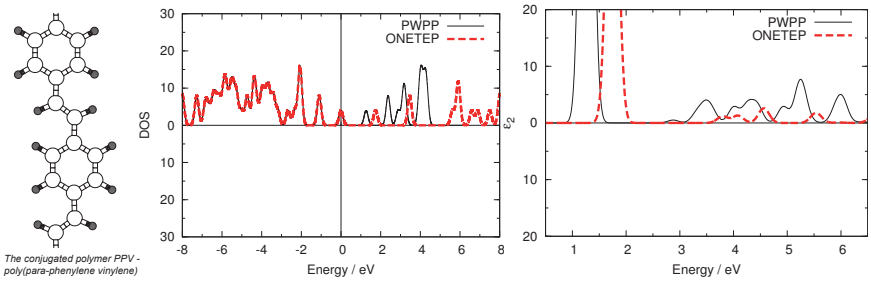
EPSRC Engineering and Physical Sciences Research Council
The Leverhulme Trust
THE WINTON PROGRAMME FOR THE Physics of Sustainability

- Funding: EPSRC, The Leverhulme Trust (ECF), Winton Programme for the Physics of Sustainability (fellowship)
- Computing: ARCHER (UK National Supercomputers), Warwick Centre for Scientific Computing, Cambridge HPC Facility (DARWIN)
- TMDC & BP/hBN calculations: G. Constantinescu, TCM, Cambridge;
- TMDC experiments: N. Wilson, Warwick, X. Xu, D, Cobden, U. Washington
- ONETEP Developers Group (P. Haynes, A. Mostofi, C. Skylaris, M. Payne, J. Dziedzic);

THE UNIVERSITY OF WARWICK

Unoccupied states and Optical Spectra

- Unoccupied states (and hence optical spectra) are not well-represented in the valence NGWF representation:

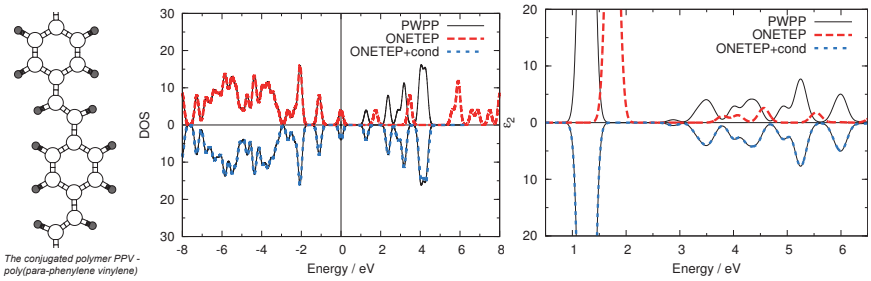


- Construct new set of NGWFs and new 'conduction state' kernel to describe ψ_c; Project out and shift valence states so that conduction states are lowermost

L. E. Ratcliff, N.D.M. Hine, P.D. Haynes, Phys. Rev. B **84**, 165131 (2012)

Unoccupied states and Optical Spectra

- Unoccupied states (and hence optical spectra) are not well-represented in the valence NGWF representation:



- Construct new set of NGWFs $\chi_\alpha(\mathbf{r})$ and new 'conduction state' kernel $K^{\alpha\beta}_{\{c\}}$ to describe ψ_c; Project out and shift valence states so that conduction states are lowermost
- Optimised conduction NGWFs can describe all localised states of a molecule (but not vacuum states)
- However, limitations of DFT mean bandgaps underestimated. TDDFT can improve this for small systems and localised excitations

L. E. Ratcliff, N.D.M. Hine, P.D. Haynes, Phys. Rev. B **84**, 165131 (2012)

Linear-Response TDDFT

Casida Formalism: Solve

$$\begin{pmatrix} \mathbf{A}(\omega) & \mathbf{B}(\omega) \\ -\mathbf{B}(\omega) & -\mathbf{A}(\omega) \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$$

where

$$\begin{aligned} A_{cv,c'v'}(\omega) &= \delta_{cc'}\delta_{vv'}(\varepsilon_{c'}^{\mathrm{KS}} - \varepsilon_{v'}^{\mathrm{KS}}) + K_{cv,c'v'}(\omega) \\ B_{cv,c'v'}(\omega) &= K_{cv,v'c'}(\omega) \end{aligned}$$

and

$$K_{cv,c'v'}(\omega) = \left\langle \psi_c^{\mathrm{KS}}\psi_v^{\mathrm{KS}} \left| \frac{1}{|\mathbf{r}-\mathbf{r}'|} + f_{\mathrm{xc}}(\mathbf{r},\mathbf{r}',\omega) \right| \psi_{c'}^{\mathrm{KS}}\psi_{v'}^{\mathrm{KS}} \right\rangle$$

Tamm-Dancoff approximation: $\mathbf{Y} = 0$; $\mathbf{B} = 0$; $\mathbf{AX} = \omega\mathbf{X}$
Use iterative eigensolvers: Only the action $\mathbf{q} = \mathbf{Ax}$ is required.
ω-dependence of f_{xc} is generally dropped (eg ALDA)

LR-TDDFT with LS-DFT

Drop ω-dependence of f_{xc} for local functionals (eg ALDA)

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B} & -\mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$$

Make Tamm-Dancoff approximation (ignore de-excitations): $\mathbf{Y} = 0$; $\mathbf{B} = 0$; Can then use iterative eigensolvers.

$$\mathbf{AX} = \omega\mathbf{X} \quad \omega = \min_{\mathbf{X}} \frac{\mathbf{X}^\dagger\mathbf{AX}}{\mathbf{X}^\dagger\mathbf{X}}$$

Write $\mathbf{q} = \mathbf{AX}$ via effective transition density $\rho^{\{1\}}(\mathbf{r})$ and KS eigenvalues

$$q_{cv} = (\varepsilon_c^{\mathrm{KS}} - \varepsilon_v^{\mathrm{KS}})X_{cv} + \left(V_{\mathrm{SCF}}^{\{1\}}\left[\rho^{\{1\}}\right]\right)_{cv}$$

where $\rho^{\{1\}}(\mathbf{r}) = \sum_{cv} \psi_c^{\mathrm{KS}}(\mathbf{r}) X_{cv}\, \psi_v^{\mathrm{KS}}(\mathbf{r})$

Express transition density via NGWFs & response density matrix $\mathbf{P}^{\{1\}}$

$$\rho^{\{1\}}(\mathbf{r}) = \sum_{\alpha\beta} \chi_\alpha(\mathbf{r}) P^{\{1\}\alpha\beta} \phi_\beta(\mathbf{r})$$

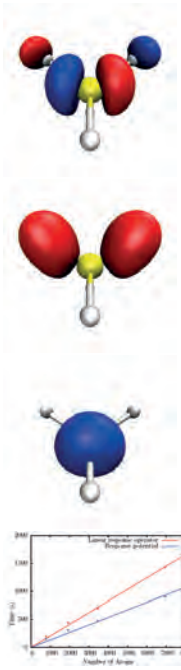
Hole / Electron each well-described by $\{\phi_\alpha(\mathbf{r})\}$ / $\{\chi_\alpha(\mathbf{r})\}$ respectively

TDDFT gradient $\mathbf{q} = \mathbf{Ax}$ in $\{\phi\}$ / $\{\chi\}$ representation:

$$\mathbf{q}^{\chi\phi} = \mathbf{P}^{\{c\}}\mathbf{H}_{\mathrm{KS}}^{\chi}\mathbf{P}^{\{1\}} - \mathbf{P}^{\{1\}}\mathbf{H}_{\mathrm{KS}}^{\phi}\mathbf{P}^{\{v\}} + \mathbf{P}^{\{c\}}\mathbf{V}_{\mathrm{SCF}}^{\{1\}\chi\phi}\mathbf{P}^{\{v\}}$$

Fully $O(N)$ for each ω if $\mathbf{P}$'s truncated.

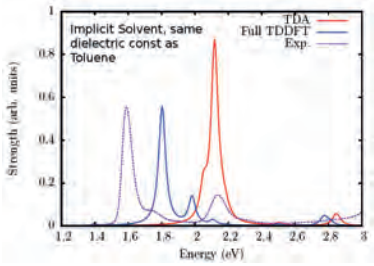
Optimise multiple $\{\mathbf{P}_i\}$ simultaneously $\rightarrow$ Scales as $O(N) \times O(N_\omega)$ with small $O(N) \times O(N_\omega^2)$ orthogonalisation step



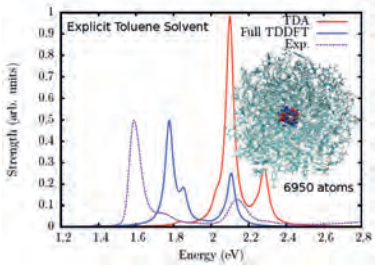
T. J. Zuehlsdorff, N. D. M. Hine, [...], P. D. Haynes , J. Chem. Phys. 139 064104 (2013).

Full TDDFT & Response Kernel Truncation

TDA often inadequate - use full TDDFT: arxiv.org/abs/1507.08157

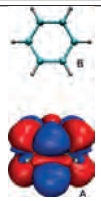


Implicit Solvent description does not give solvatochromic shift, use explicit solvent:



0.5eV error down to 0.17eV & nearly green!

Truncation of response kernel can be used to eliminate spurious charge-transfer states

Fully dense $P^{\{1\}}$	**$P^{\{1\}}$ on Benzene A**
10th excitation:	1st excitation:
$\omega = 5.1950$ eV	$\omega = 5.1953$ eV
$f = 0.177 \times 10^{-6}$	$f = 0.111 \times 10^{-6}$
	

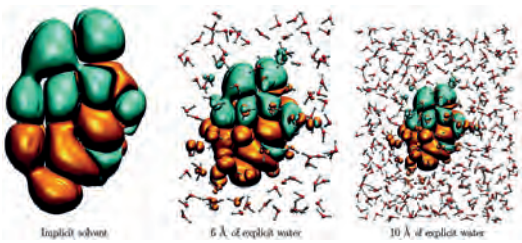
Atom-centred NGWFs provide natural means to constrain excitations to subspaces

$\Rightarrow$ significant increase in efficiency

Can reintroduce subsystem coupling perturbatively with final subspace diagonalisation: $\mathbf{P}^{\{1\}}_{\text{tot}} = \sum_I \alpha_I \mathbf{P}^{\{1\}}_{\mathbf{A}_I} + \sum_j \beta_J \mathbf{P}^{\{1\}}_{\mathbf{B}_J}$

Alizarin: Solvatochromic shifts

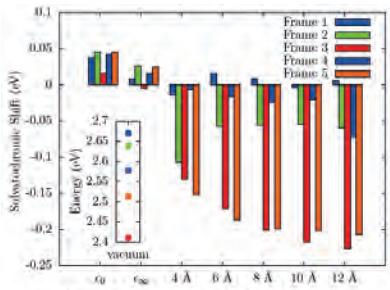
Alizarin prototypical red dye exhibiting strong solvatochromic shift



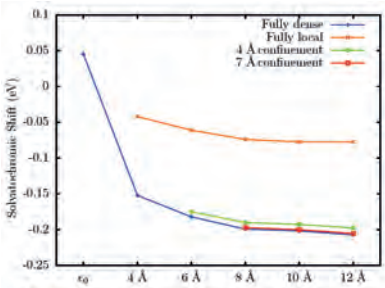
S_1 excitation stays mostly localised but fractions of electron and hole delocalise to water $\Rightarrow$ slow convergence of solvatochromic shift

Two contributions:

- Electrostatic potential of specific configuration
- Delocalisation of transition to environment



Convergence of S_1 with explicit water



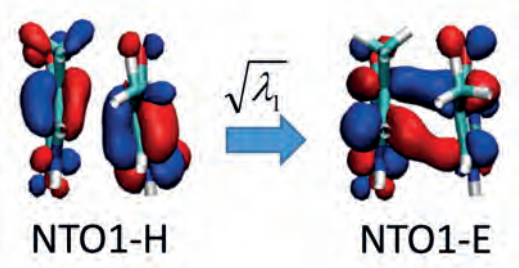
Convergence with response density matrix truncation

Quantified Natural Transition Orbital Analysis

Large systems → many low-lying excitations

How to follow excitation as geometry varies?

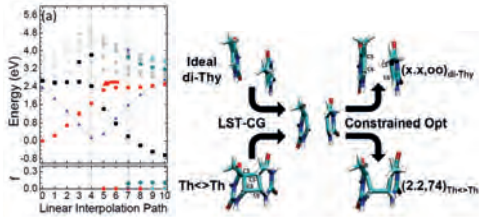
QNTO analysis allows adiabatic PES to be "reconnected" by matching transition origins.



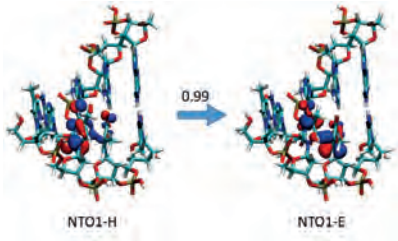
Find **U**, **V** matrices that diagonalise transition density matrix.

$$\mathbf{P}_I^{\{1\}} = \bar{\mathbf{U}}_I \mathbf{D}_I \bar{\mathbf{V}}_I$$

Identify states involved in funnel resulting in thymine dimerisation



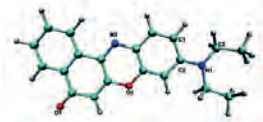
Mechanism then verified in much larger model (4 base pairs of DNA)



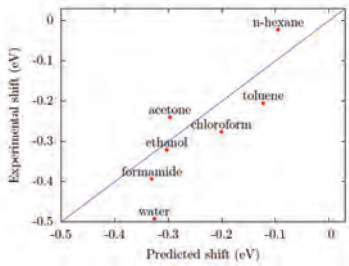
J.-H. Li, T. J. Zuehlsdorff, M. C. Payne and N. D. M. Hine, Phys. Chem. Chem. Phys., 17, 12065 (2015)
J.-H. Li, T. J. Zuehlsdorff, M. C. Payne and N. D. M. Hine, upcoming (2015)

Colour Prediction

Nile Red (9-diethylamino-5-benzo[α]phenoxazinone)



Strong solvent-dependent shift:



Not well-predicted by electrostatic models

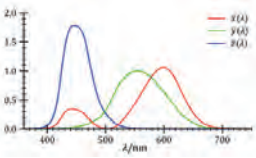
Can we use the absorption spectrum to directly predict the colour?

$$I_{\text{trans}} = I_0(\lambda) e^{-\kappa(\lambda)x}$$

$\kappa(\lambda)$: absorption coefficient;
$I_0(\lambda)$: spectrum of illuminant;

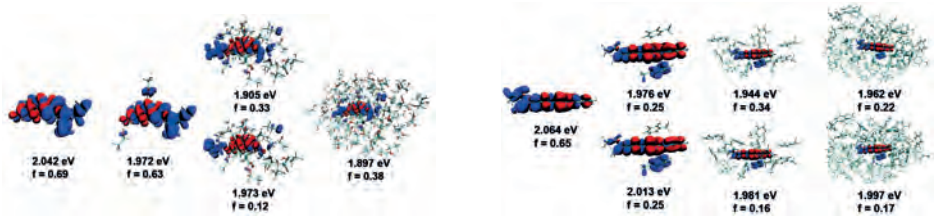
$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = N \int I_{\text{trans}}(\lambda) \begin{pmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{pmatrix} \mathrm{d}\lambda$$

$\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$: Tristimulus color matching functions

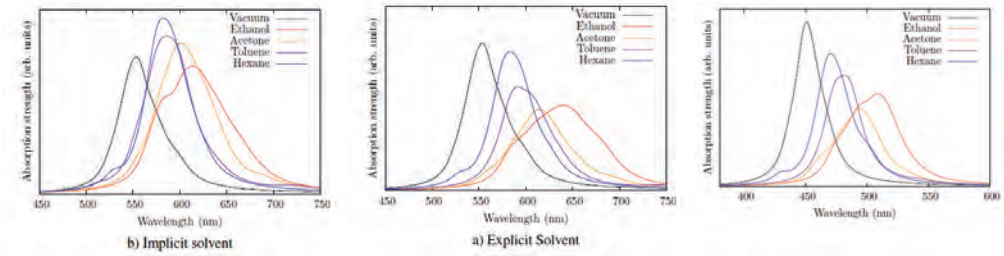


X, Y, Z can be transformed into RGB.

Nile Red Colour Prediction



Ethanol: hydrogen bonded configurations; Toluene: $\pi - \pi$ stacked configuration.
Run long explicit solvent MD - extract samples

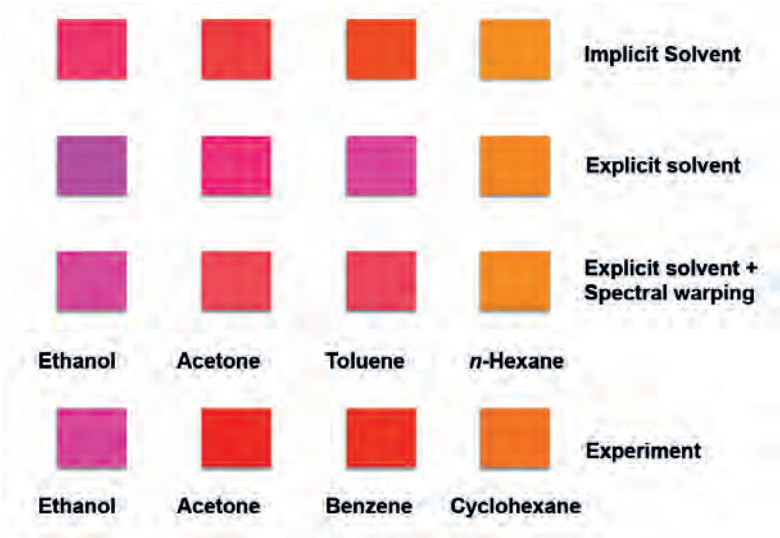


Spectral warping: calculate CAM-B3LYP spectrum, apply transformation to PBE spectrum:

$$\omega_i^{\text{trans}} = \omega_i + \beta + \alpha\phi_i^2$$

Nile Red Colour Prediction

Sufficient accuracy for useful predictive power:



Not yet better than colour-resolving power of the eye, but pretty good!

Mantid for Data Analysis

Nick Draper (Tessella)

MANTID

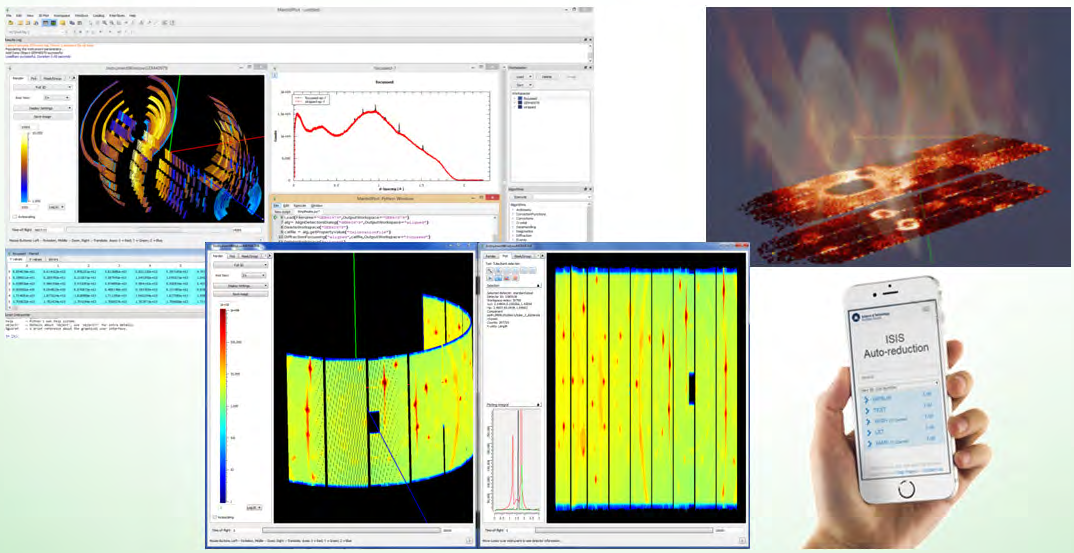
EUROPEAN SPALLATION SOURCE

Science & Technology Facilities Council ISIS

OAK RIDGE National Laboratory
HIGH FLUX ISOTOPE REACTOR
SPALLATION NEUTRON SOURCE

Tessella
Technology & Consulting

MANTID



THE MANTID PROJECT

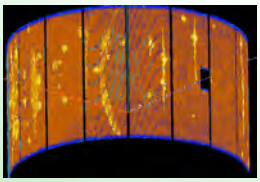
Project Goals

- Goals
  - Consolidate the data reduction/analysis software for neutron scattering experiments
  - Without restricting the needs of the instrument scientists
- Key requirements
  - Cross-platform
  - Easily extensible
  - Freely redistributable
  - Open source

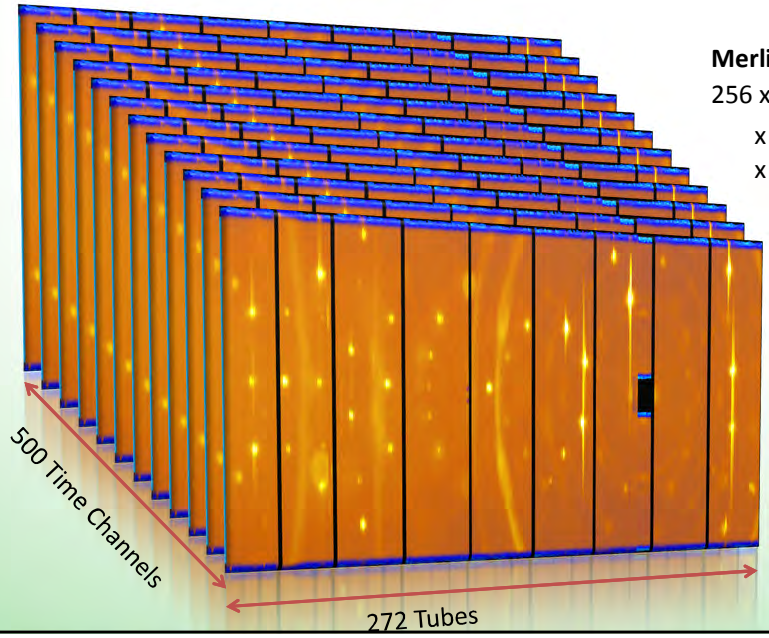
Scientific Data Flow

Raw Data

- Time & Detector Coords
- Full of measurement artefacts
- Useful for Diagnostics & Fault finding



Raw Data



Merlin

256 x 272 = 70K Pixels

x 500 = 35M Pixels

x 5 incident energies
= 175M Pixels

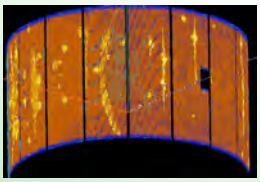
Scientific Data Flow

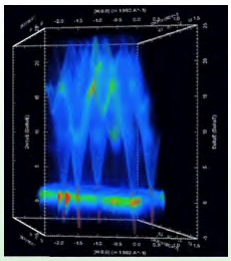
Raw Data

- Time & Detector Coords
- Full of measurement artefacts
- Useful for Diagnostics & Fault finding

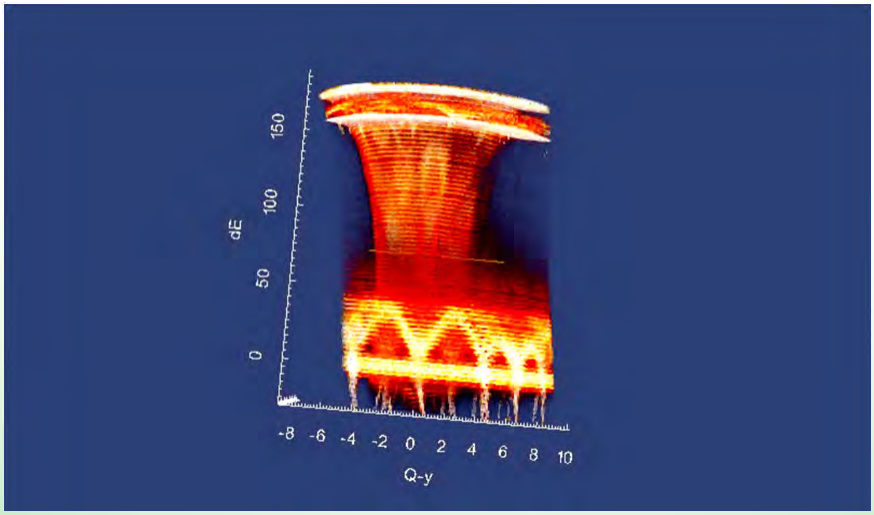
Reduced Data

- Measurement artefacts corrected
- Scientific Units
- Size: compressed or expanded





Multi Dimensional Visualization

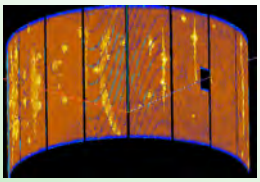


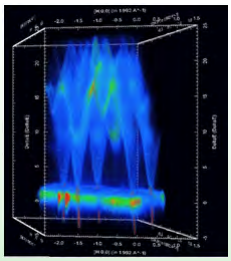
Scientific Data Flow

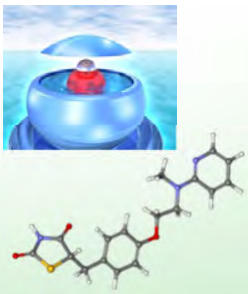
Raw Data
- Time & Detector Coords
- Full of measurement artefacts
- Useful for Diagnostics & Fault finding

Reduced Data
- Measurement artefacts corrected
- Scientific Units
- Size: compressed or expanded

Data Analysis
- Comparison with physical models
- Extraction of understanding







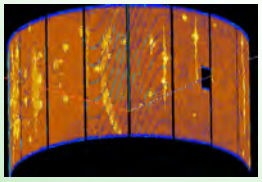
Scientific Data Flow

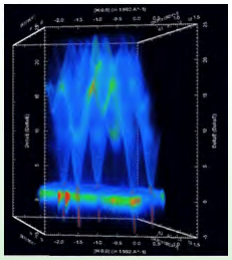
Raw Data
- Time & Detector Coords
- Full of measurement artefacts
- Useful for Diagnostics & Fault finding

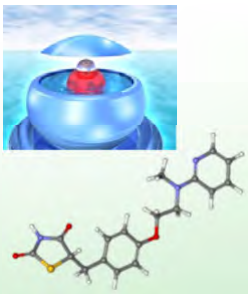
Reduced Data
- Measurement artefacts corrected
- Scientific Units
- Size: compressed or expanded

Data Analysis
- Comparison with physical models
- Extraction of understanding

Impact
- Scientific Publication
- Commercialization
- Process changes
- New materials







Scientific Data Flow

Raw Data
- Time & Detector Coords
- Full of measurement artefacts
- Useful for Diagnostics & Fault finding

Reduced Data
- Measurement artefacts corrected
- Scientific Units
- Size: compressed or expanded

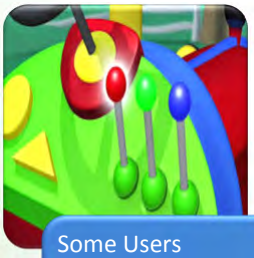
Data Analysis
- Comparison with physical models
- Extraction of understanding

Impact
- Scientific Publication
- Commercialization
- Process changes
- New materials

- Visualization
- Algorithms
- Model Fitting & Optimisation
- Live Data

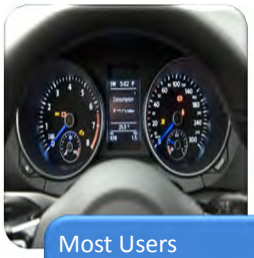
MANTID

Matching the users experience



Some Users

- Technique specific interfaces
- Point and click interface
- Simple Graphing
- Detailed help
- Training & Examples



Most Users

- Flexible user interface
- Python Scripting
- Data History
- Powerful visualisations

Super Users

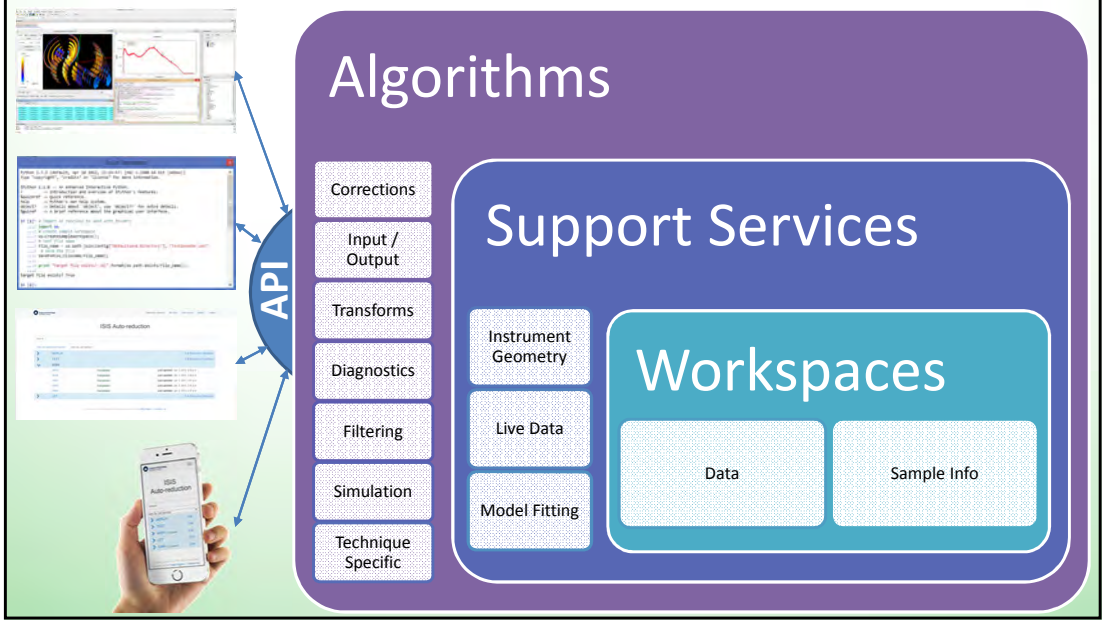
- Plug ins
- Python Algorithms
- Custom User Interfaces

MANTID

Professional development

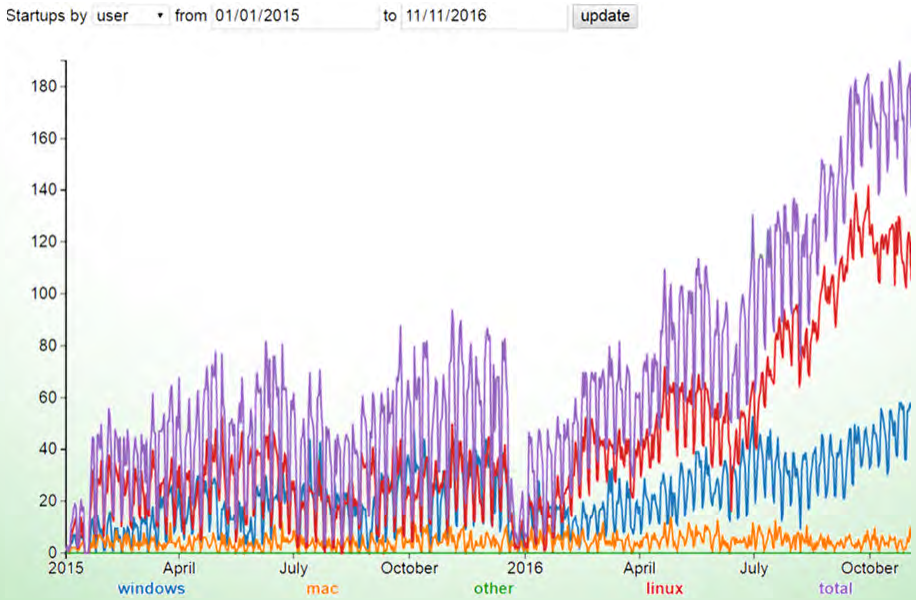
- Distributed Development Team
  - International Project Governance Board & Scientific Steering Committee
  - Dedicated software engineers
    - All with science BSc, many with PhD
  - Working closely with Instrument Scientists
- Build and testing automation
  - 20+ build servers
  - World class static code analysis
  - 11,900 unit tests
  - 350 automated system tests
- Comprehensive documentation
  - 868 automated tests on included scripts
  - Both online and installed with Mantid

Architecture



International Collaboration

Usage



http://reports.mantidpoject.org

Analysis In Mantid

Reduction vs Analysis

- Mantid's key scope is Data reduction
- Some techniques have accepted tools for analysis
  - GSAS for diffraction
  - SASView for SANS
- Then Mantid interfaces with the accepted tools
- Others don't
  - Here Mantid goes further itself

Fitting

- Simple and comprehensive fitting
  - Visual or scriptable
  - Build simple or complex models
  - Constraints
  - Choice of Minimizers
  - Single or Multiple spectra
  - Extensible
    - Math
    - Python
    - C++
  - 1D, 2D -> nD model fitting

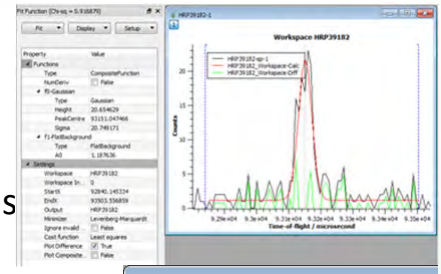


Table-1 - HRP39182_Parameters

	Name[L]	Value[Y]	Error[yEr]
0	f0.Height	20.6546	2.50495
1	f0.PeakCentre	93151	2.27126
2	f0.Sigma	20.7492	1.82465
3	f1.A0	1.18764	0.157508
4	Cost function value	0.916879	0

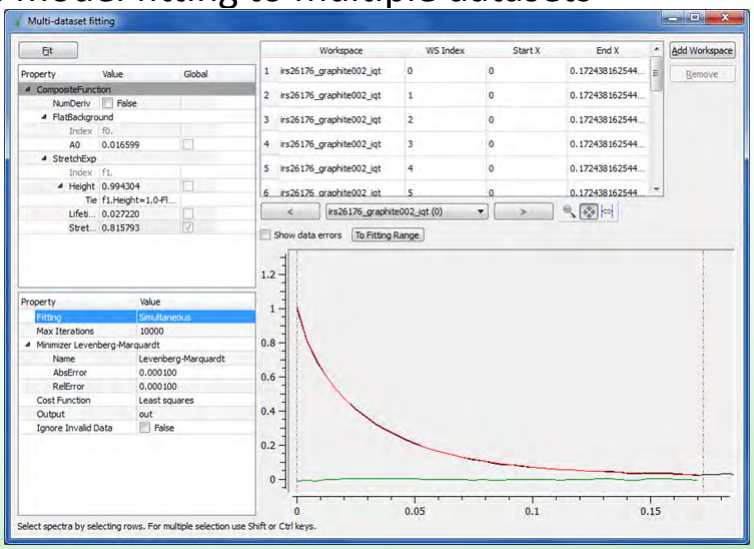
HRP39182_Workspace - Mantid

Y values | X values | Errors

	0	1	2
Data	1.000000e+00	0.000000e+00	2.000000e+00
Calc	1.187636e+00	1.187636e+00	1.187636e+00
Diff	-1.876357e-01	-1.187636e+00	8.123643e-01
Gaussian	0.000000e+00	0.000000e+00	0.000000e+00
FlatBackground	1.187636e+00	1.187636e+00	1.187636e+00

Multi dataset fitting

- Simultaneous model fitting to multiple datasets
- Same model building and constraints
- Local or Global parameters



Fit functions as PLUGINS

- Mantid's fitting functions can be added to by user-defined plugins
- Written in Python
- Optimization framework oblivious to plugin nature
  - Free to use any minimizer, cost function available in Mantid

Simple Example: LINEAR

- Place the following in a file called 'LinearFunction.py' on Mantid's python plugin path (see MantidPlot preferences)

```
from mantid.api import IFunction1D, FunctionFactory

class LinearFunction(IFunction1D):
    def init(self):
        self.declareParameter("A", 1.0)
        self.declareParameter("B", 1.0)

    def function1D(self, x):
        # x is a 1D numpy array
        a = self.getParameterValue("A")
        b = self.getParameterValue("B")
        # Mantid expects a 1D numpy array the same size as x
        return a*x + b

# Register with Mantid
FunctionFactory.subscribe(LinearFunction)
```

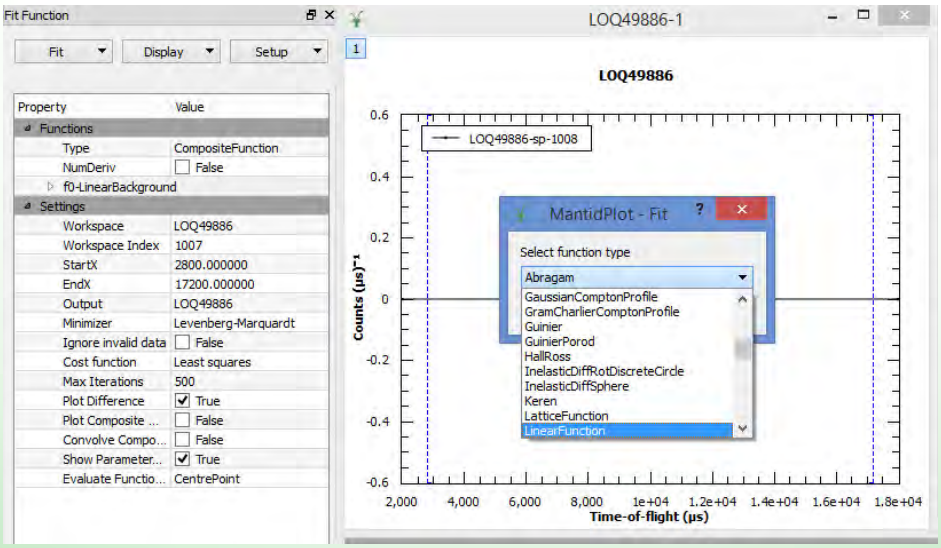
Adding Derivatives

- Analytic derivatives can be added if desired. In most scenarios the numerical ones should be adequate

```
from mantid.api import IFunction1D, FunctionFactory

class LinearWithDerivative(IFunction1D):

    # … snipped. Same as previous slide …

    def functionDeriv1D(self, x, jacobian):
        # Compute the matrix of partial derivatives
        for i, x_i in enumerate(x):
          jacobian.set(i, 0, 1) # paramter at index 0
          jacobian.set(i, 1, x_i) # paramter at index 1


# Register with Mantid
FunctionFactory.subscribe(LinearWithDerivative)
```

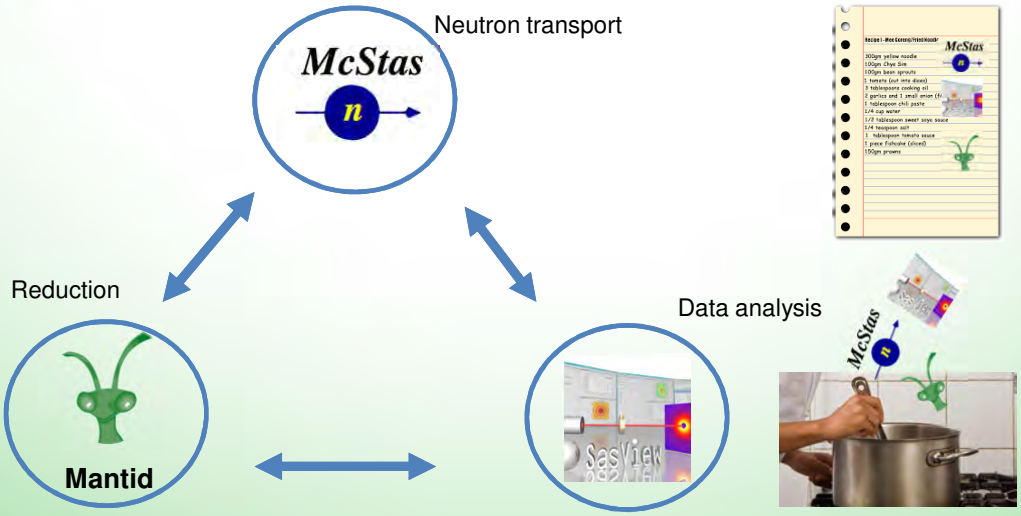
- For more information see our final training course:
  - www.mantidproject.org/Extending_Mantid_With_Python

Available in Mantidplot



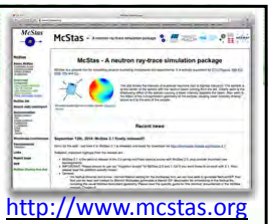
Linking mantid, mcstas and sasview

"Full-circle" virtual experiment software integration



Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

McStas Introduction



http://www.mcstas.org

- Flexible, general simulation utility for neutron scattering experiments.
- Original design for Monte carlo Simulation of triple axis spectrometers
- Developed at DTU Physics, ILL, PSI, Uni CPH, **ESS**
- V. 1.0 by K Nielsen & K Lefmann (1998) RISØ
- Currently 2.5+1 people full time plus students

GNU GPL license
Open Source

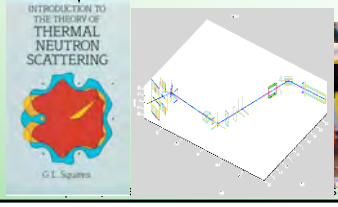
Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

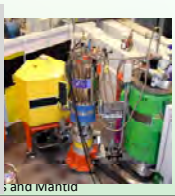
What is McStas used for?

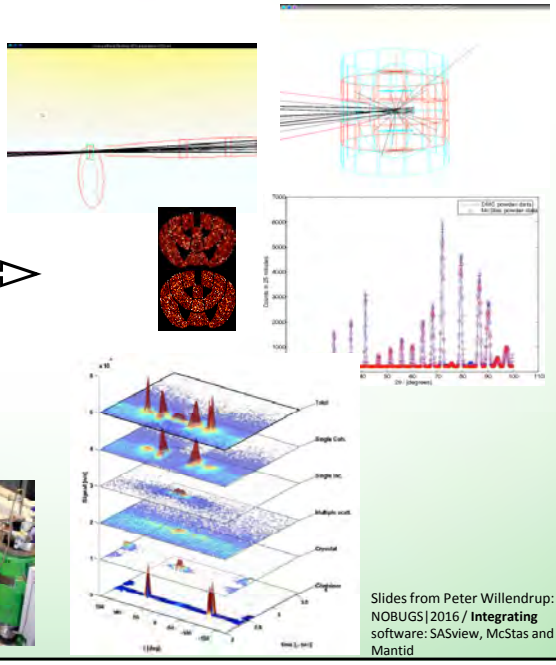
McStas

- Instrumentation
- Virtual experiments
- Data analysis
- Teaching

(KU, DTU, e-Learning…)







Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

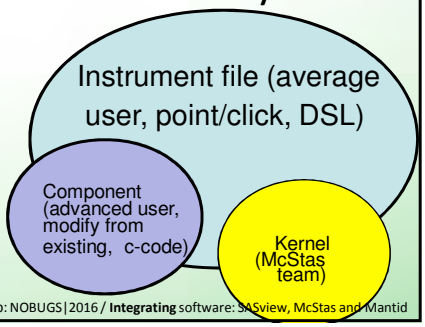
McStas overview

- Portable code (Unix/Linux/Mac/Windows)

- Ran on everything from iPhone to 1000+ node cluster!

- 'Component' files (~150) inserted from library
  - Sources
  - Optics
  - Samples
  - Monitors
  - If needed, write your own comps (**short ~200 LoC**)

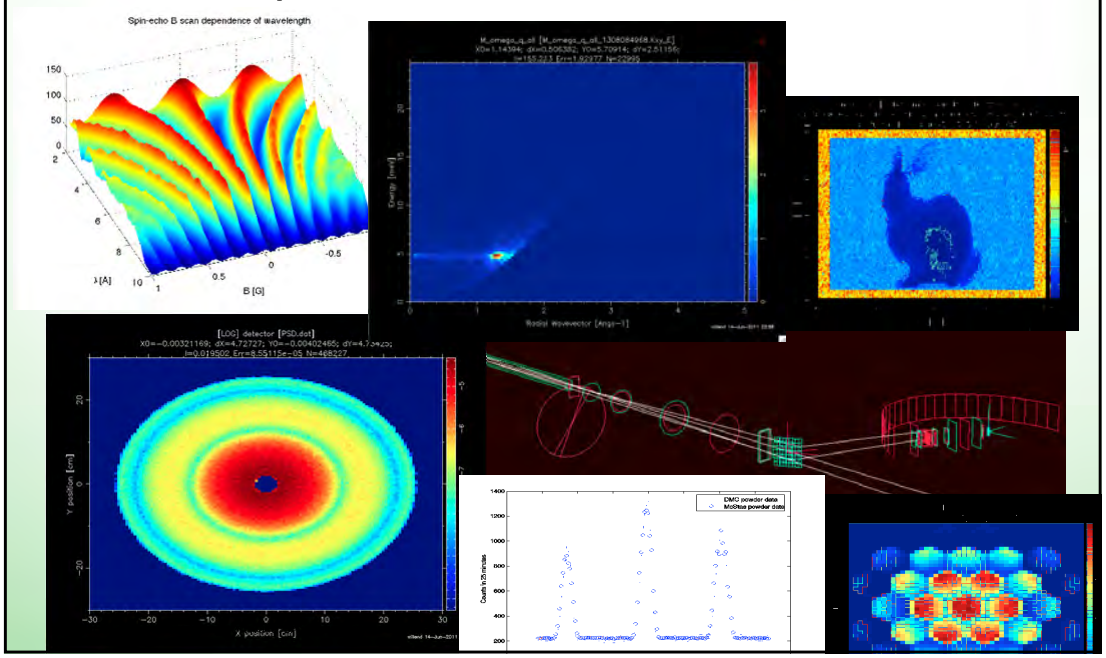
- DSL + **ISO-C** code gen.



Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

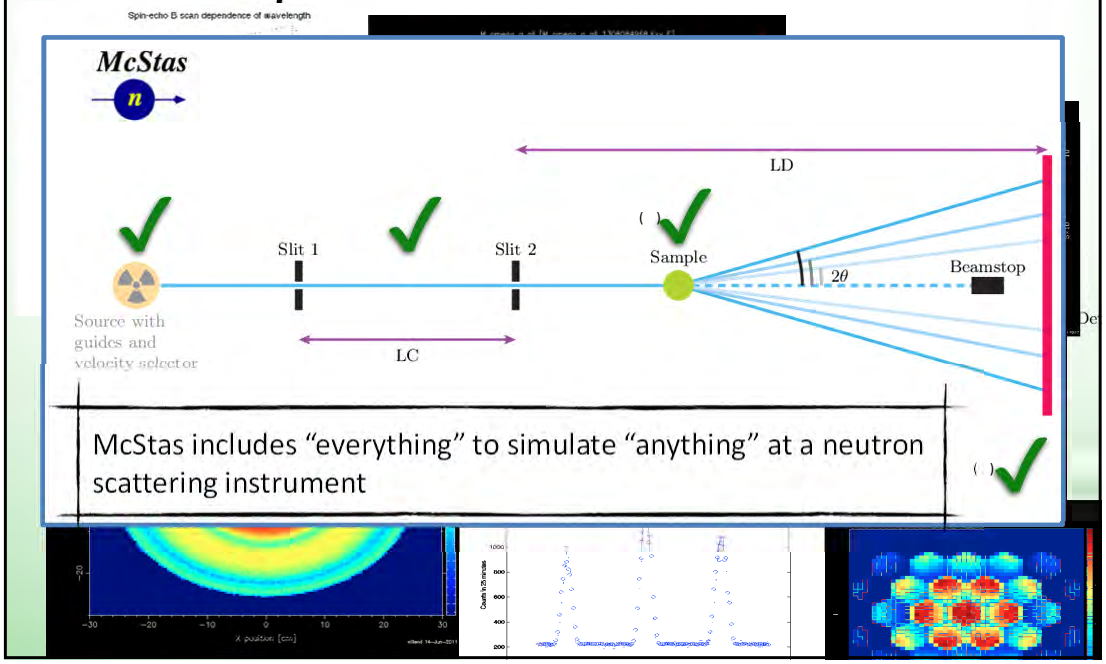
MANTID

Example suite: 123 instruments



MANTID

Example suite: 123 instruments



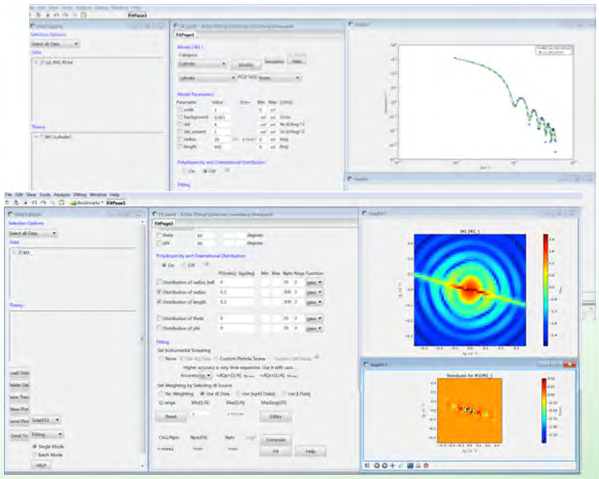
McStas includes “everything” to simulate “anything” at a neutron scattering instrument

SasView - Small Angle Scattering Analysis

- Operates on reduced scattering data
- **Performs modeling in inverse space**
- Data analysis toolbox:
  - Fitting **models** to data
  - P(r) inversion
  - Model-independent analysis
- Other useful tools

SasView - Fitting

- Handles 1D and 2D data
- support for canSAS XML & NXcanSAS formats
- **Form and structure factors for various particle shapes**
- Different optimizers (Bayesian Statistics)
- Allows polydispersity
- Simultaneous and batch fitting



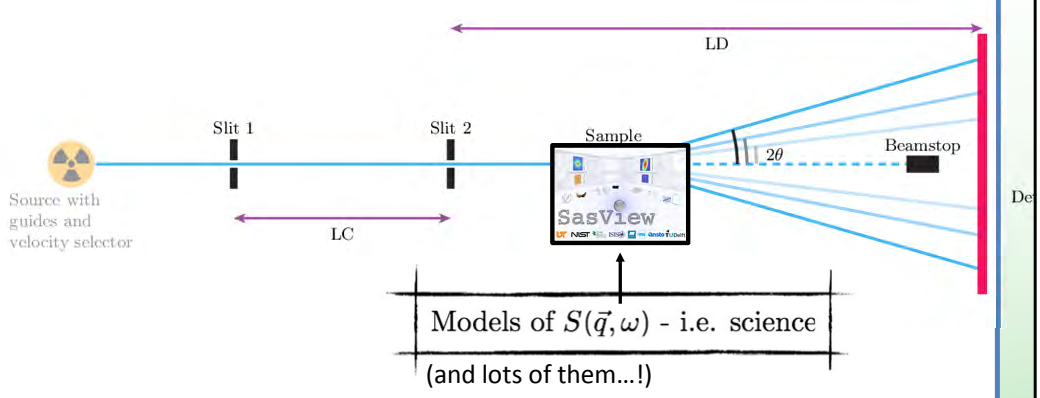
SasView 4.0 is out

- SasView "built-in" models have been separated out into an **independent package (aka SASmodels)**
- Easy to add custom user models (including advanced)
- Support for OpenCL
- All model documentation has been reviewed and updated
- Number of minor bugs fixed

Available from:
https://github.com/SasView/sasview/releases/tag/v4.0

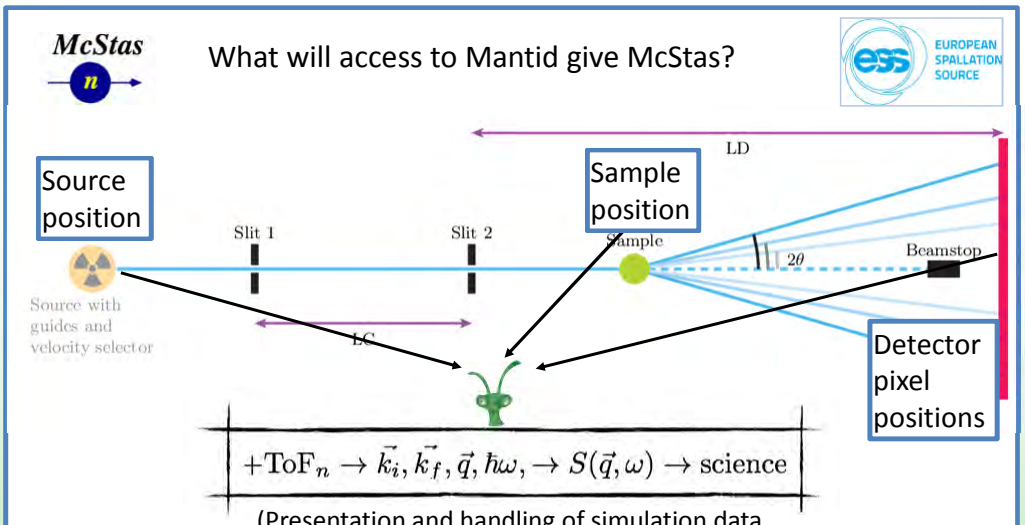
SasView 4.0 is out

What will access to SASview give McStas?



MANTID

What will access to Mantid give McStas?



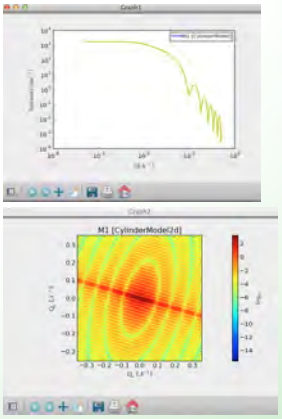
(Presentation and handling of simulation data ala non-virtual ToF instruments…)

Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

MANTID

Why – instrument development

- ✓ McStas used for instrument design work at
- ✓ LoKI team needed 2D scattering kernels in McStas
- ✓ Use SasView scattering kernels in McStas



SasView scattering kernel: Orientated cylinder

- ✓ Mantid is reduction software for
- Make solution with all 3 for bootstrapping the analysis pipeline!

Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

McStas mcdoc: SasView_model

[Identification | Description | Input parameters | Output parameters | Links]

The `SasView_model` Component

This SANS sample exposes SasView's scattering kernels to McStas. In this way SasView's monodisperse scattering kernels can be call from McStas.

Identification

- **Author:** Jakob Garde, Torben Nielsen, Peter Willendrup
- **Origin:** SasView, DTU, European Spallation Source ERIC
- **Date:** 03.02.2016
- **Version:** 0.5

Description

```
Sample for use in SANS instruments. The models describe mono disperse particles in thin solution. The sample geometry may have the shape:

Shape:          - A filled box with dimensions xwidth, yheight and zdepth.
                - A cylinder with dimensions radius and yheight.
                - A filled sphere given by radius.

These parameters are mutually exclusive.

Example using spheres in thin solution, with radius=200 AA and a delta_sld=0.6 fm/AA^3:
     SasView_model(model_index=47, model_scale=1.0, model_pars={1, 7, 200}, model_abs=0.0,
                   xwidth=0.01, yheight=0.01, zdepth=0.005,)


The algorithm of this component requires use of the ISO C standard c99. You could use MCSTAS_CFLAGS like MCSTAS_CFLAGS = -std=c99 -g -O2 -lm.
Or, add the flag "-std=c99" to "C flags" in the mcgui or mcgui-py config.

The list of scattering models in SasView is called sasmodels. The list of McStas available SasView sasmodels are found in the table below.

A few models may require manual documentation lookup using the above link to the SasView site.

McStas does currently not support multiplication of formfactor models with structure factor models.

The 2D scattering scattering kernels are denoted by modelname_xy. I.e. to evalulate scattering from aligned cylinders use model_index=10 to use cylinder_xy.
```

Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

McStas mcdoc: SasView_model

Model no.	SasView name	Parameters
0	None	None
1	barbell	(sld, solvent_sld, bell_radius, radius, length)
2	barbell_xy	(sld, solvent_sld, bell_radius, radius, length, theta, phi)
3	bcc_paracrystal	(dnn, d_factor, radius, sld, solvent_sld)
4	bcc_paracrystal_xy	(dnn, d_factor, radius, sld, solvent_sld, theta, phi, psi)
5	capped_cylinder	(sld, solvent_sld, radius, cap_radius, length)
6	capped_cylinder_xy	(sld, solvent_sld, radius, cap_radius, length, theta, phi)
7	core_shell_cylinder	(core_sld, shell_sld, solvent_sld, radius, thickness, length)
8	core_shell_cylinder_xy	(core_sld, shell_sld, solvent_sld, radius, thickness, length, theta, phi
9	cylinder	(sld, solvent_sld, radius, length)
10	cylinder_xy	(sld, solvent_sld, radius, length, theta, phi)
...	...	...
58	triaxial_ellipsoid_xy	(sld, solvent_sld, req_minor, req_major, rpolar, theta, phi, psi)

McStas

Normal McStas interface

...

Available SasView sasmodels - Extended parameters

sasmodels name	Parameters	Units	Range
barbell	{ 4, 1, 40, 20, 400 }		
barbell_xy	{ 4, 1, 40, 20, 400, 60, 60 }		
	Parameters:		
	Barbell scattering length density	4e-6/Ang^2	[-inf, inf]
	Solvent scattering length density	1e-6/Ang^2	[-inf, inf]
	Spherical bell radius	Ang	[0, inf]
	Cylindrical bar radius	Ang	[0, inf]
	Cylinder bar length	Ang	[0, inf]
	In plane angle	degrees	[-inf, inf]
	Out of plane angle	degrees	[-inf, inf]

...

with access to SASview models and their docs

...

McStas mcdoc: SasView_model

- Links to SasView documentation

2.1.14. CylinderModel

This model provides the form factor for a right circular cylinder with uniform scattering length density. The form factor is normalized by the particle volume.

For information about polarised and magnetic scattering, click here.

2.1.14.1. Definition

The output of the 2D scattering intensity function for oriented cylinders is given by (Guinier, 1955)

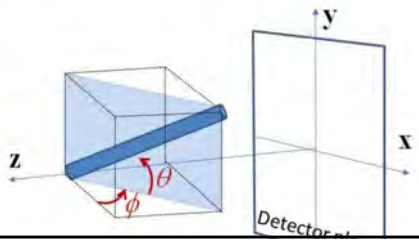
$$P(q,\alpha) = \frac{scale}{V} f^2(q) + bkg$$

where

$$f(q) = 2(\Delta\rho)V \sin(qL\cos\alpha/2)/(qL\cos\alpha/2)\frac{J_1(qr\sin\alpha)}{(qr\sin\alpha)}$$

and α is the angle between the axis of the cylinder and the *q*-vector, *V* is the volume of the cylinder, *L* is the length of the cylinder, *r* is the radius of the cylinder, and Δρ (contrast) is the scattering length density difference between the scatterer and the solvent. *J1* is the first order Bessel function.

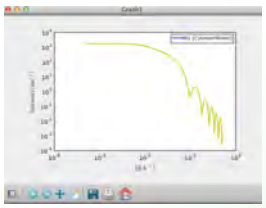
To provide easy access to the orientation of the cylinder, we define the axis of the cylinder using two angles θ and φ. Those angles are defined in Figure 1.



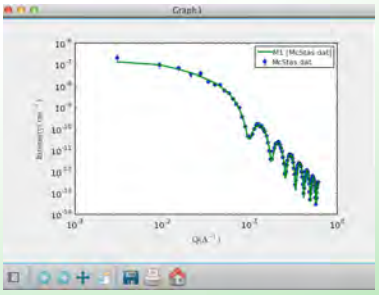
Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

SasView -> McStas -> SasView

- SasView cylinder model (1D) scattering kernel used in McStas
- radius = 40 Å, length = 400 Å



- McStas event data saved as I(q)
- SasView reads McStas I(q) data



- SasView fit engine gives:
- radius = 40 Å
- Length = 401 Å

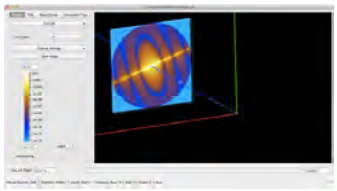
Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

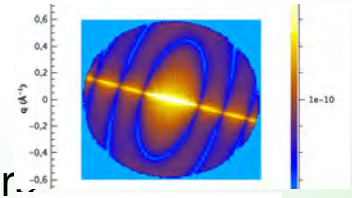
SasView -> McStas -> Mantid -> SasView

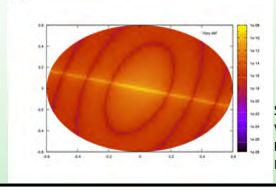
McStas -> event data

Mantid -> I(qx,qy)

SasView -> model parameters



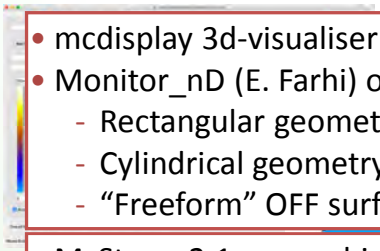




2D scattering kernel: Orientated cylinder

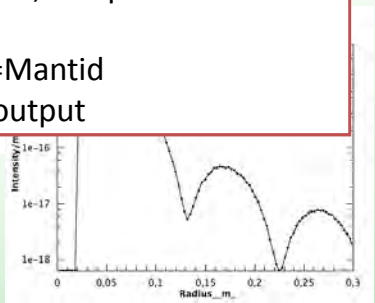
Slides from Peter Willendrup: NOBUGS|2016 / **Integrating** software: SASview, McStas and Mantid

McStas data in Mantid



- mcdisplay 3d-visualiser modified to generate Mantid IDF xml file
- Monitor_nD (E. Farhi) output event data with "Pixel ID" and ToF
  - Rectangular geometry
  - Cylindrical geometry
  - "Freeform" OFF surface geometry

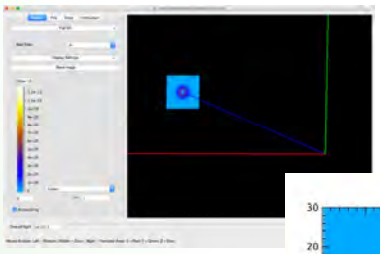
- McStas > 2.1 on machine, NeXus libs, Mantid
- Special component naming, defines source, sample and detectors for Mantid geometry
- Generate XML using mcdisplay --format=Mantid
- Run with mcrun/mcgui selecting NeXus output

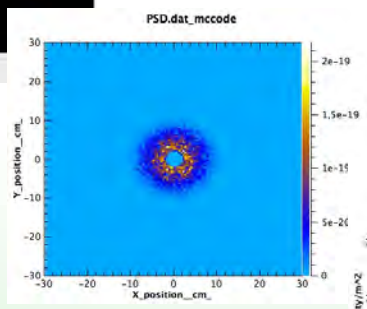


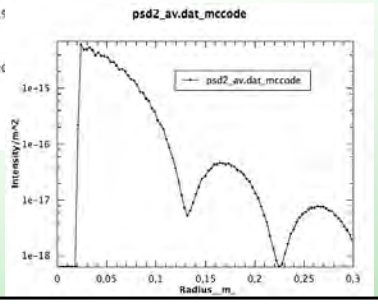
Nielsen, T.R.; Markvardsen, A.J.; Willendrup, P.: *McStas and Mantid integration*
Journal of Neutron Research, vol. 18, no. 2-3, pp. 61-77, 2015
DOI: 10.3233/JNR-16002

McStas data in Mantid

templateSANS_Mantid.instr



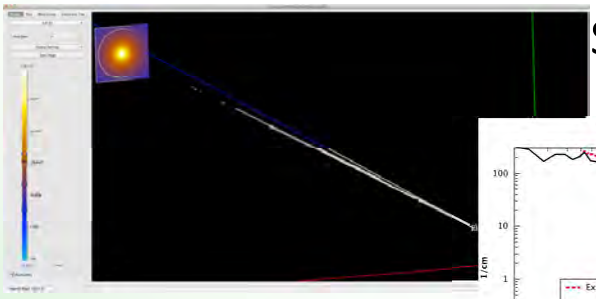


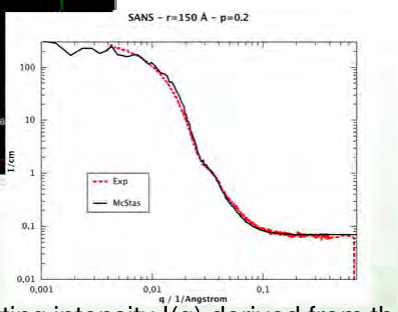


Data reduction and analysis

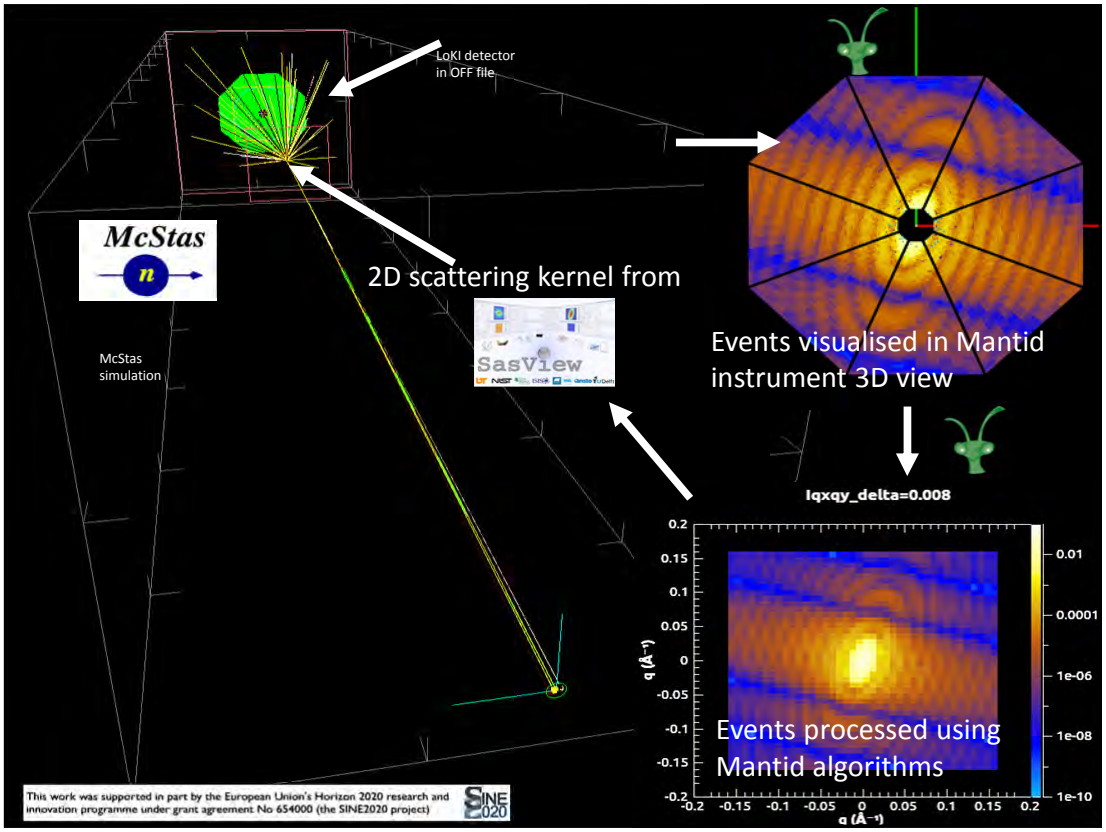
Mantid view of the McStas event data

S2d.instr file.





Comparison of rescaled scatting intensity I(q) derived from the experimental data and a McStas simulation.



MANTID

Optimization Of Molecular Dynamics Simulations

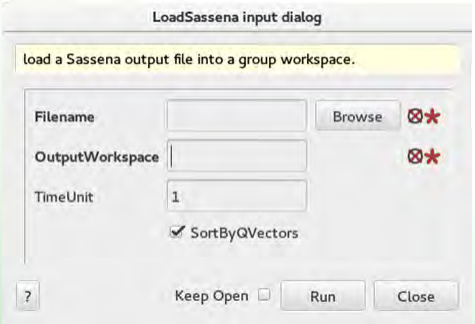
MANTID

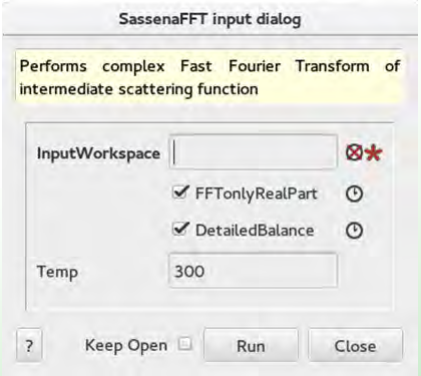
Sassena support in Mantid

- Like nMoldyn, Sassena calculates neutron and X-ray scattering functions from molecular dynamics atomic trajectories.
  → www.sassena.org

- Mantid algorithms load Sassena output and carry out its Fourier transform

$I(Q,t)$ → $S(Q,E)$



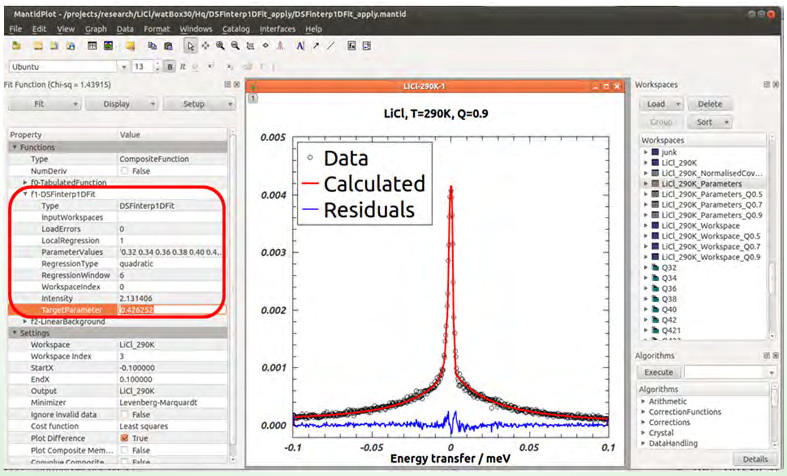


- Algorithms released as part of Mantid: www.mantidproject.org

MANTID

DSFinterp1DFit

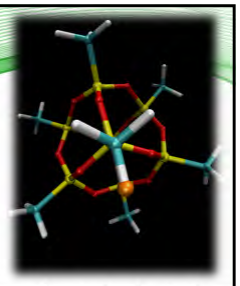
- A Mantid fitting function to optimize simulated QENS spectra against experiment.



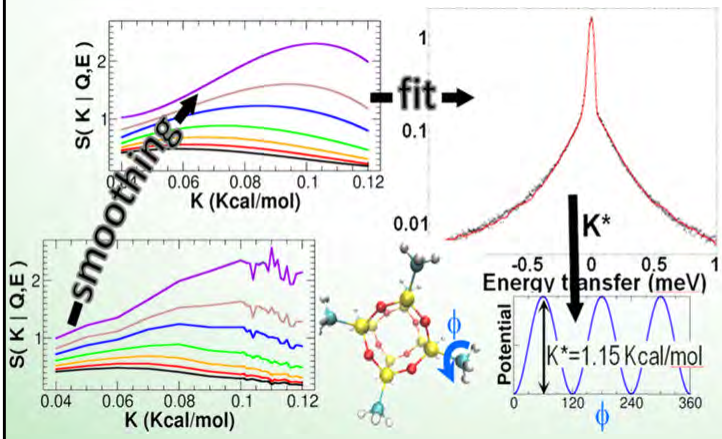
- fit function released as part of Mantid: www.mantidproject.org
- requires external python library dsfinterp (https://pypi.python.org/pypi/dsfinterp/0.1)

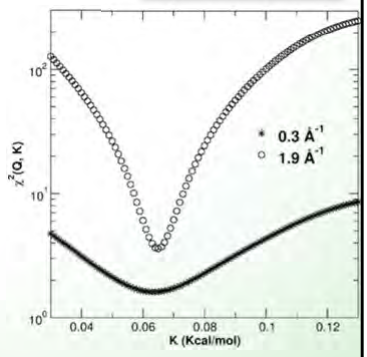
MANTID

Application of DSFinterp1DFit



- Example: optimization of the energy barrier to methyl rotations in the octamethyl silsesquioxane molecule.

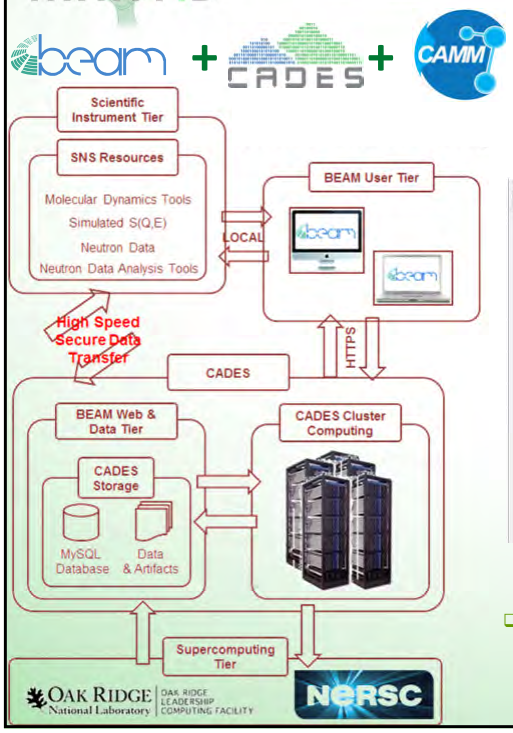




- Data taken at BASIS
- Data reduced with Mantid
- Analysis using NAMD, AmberTools, Sassena, and Mantid

Borreguero, Jose M., and Vickie E. Lynch. "Molecular dynamics force-field refinement against quasi-elastic neutron scattering data." *Journal of Chemical Theory and Computation* **12**.9 (2016).

MANTID



The Bellerophon Environment integrates **Mantid** and **Sassena**, utilizing OLCF, NERSC, and CADES compute resources to automate workflows for parameter-refinement of force-fields used in molecular dynamics simulations, by iterative optimization against QENS data

- E. Lingerfelt *et al.*, "Near Real-time Scalable Analysis of High-dimensional Nanophase Materials Imaging and Neutron Science Data in the DOE HPC Cloud with BEAM", 2015 International Conference for High Performance Computing, Networking, Storage and Analysis, Nov. 18, Austin, TX

MANTID

Conclusions

- Mantid is looking to work with, not replace accepted and loved analysis tools
- Mantid is a capable tool for comparing experimental data to simulations
- Mantid is capable of simple to complex model optimization
- With a bit more hardware you can optimize the simulations to the experimental data

MANTID

Contributors

Partners

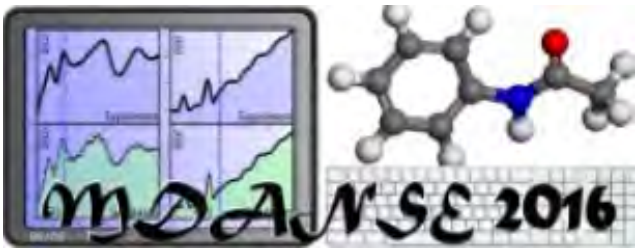
ISIS, EUROPEAN SPALLATION SOURCE, NEUTRONS FOR SCIENCE, OAK RIDGE National Laboratory, HIGH FLUX ISOTOPE REACTOR, SPALLATION NEUTRON SOURCE

Prime Contractor

Tessella Technology & Consulting

Contributors

PAUL SCHERRER INSTITUT PSI, ansto, JÜLICH FORSCHUNGSZENTRUM, McStas



… at least over the weekend …

MDANSE School 2016

Time	Thursday, 10th November	Friday, 11th November	Saturday, 12th November
9.00		Introduction to software to be used in tutorials Miguel Gonzalez, Keith Refson	Advanced Analysis in Computer Simulations (sMoldyn) Gerald Kneller
9.30		Hands on tutorials	
9.45		Two groups will work simultaneously 1.Molecular Dynamics and QENS 2. Lattice Dynamics and INS Jeff/Miguel/Sanghamitra/Keith	Path Integral MD (PIMD) Matt Probert
10.00			
10.30			Refreshments
11.00	Registrations	Refreshments	Linear scaling DFT and Applications (OneTep) Nicholas Hine
		Hands on tutorials (Continued)	
11.45			Mantid for Data Analysis Nick Draper
12.00	Lunch		
12.30		Group Photo	Wrapping up Felix Fernandez-Alonso
13.00		Lunch	Lunch
13.45	Welcome Felix Fernandez-Alonso		
14.00	Neutron Scattering & Computer Simulations Stewart Parker	Hands on tutorials (Continued)	
14.45	Basic MD and Force fields Miguel Gonzalez		
15.00			
15.30	Refreshments	Refreshments	
16.00	Introduction to Electronic Structure Sanghamitra Mukhopadhyay	Hands on tutorials (Continued)	
16.30	Introduction to Lattice Dynamics Keith Refson	Round Table Discussions	
17.00		Making Way to Bus	
17.15		ISIS Tour	
17.30	Setting up Computers for the next day		
18.00	Break		
19.00	Dinner (Crown & Thistle)	Workshop Dinner (Cosener's House)	
20.00			

You've been framed!

Don't let it go cold!

Practice, practice, practice …

A MDANSE Prize?

What Next?

TESCO

Every little helps

Some thoughts from yesterday

In the absence of infinite resources ... need to think a bit harder.

- *Joint efforts to reach critical mass, engagement from all.*
- *Coaching & training, at various levels – introductory, advanced.*
- *A new forum for discussion & exchange.*
- *The in-silico instrument scientist – some progress has been made.*
- *Databases & repositories – extension of INS database at ISIS?*
- *Get on with it - joint projects & demonstrators.*

Demonstrate progress , and document it!

thank you!

&

safe trip back

MANTID
Project Goals
• Goals
– Consolidate the data reduction/analysis software for neutron scattering experiments
– Without restricting the needs of the instrument scientists
• Key requirements
– Cross-platform
– Easily extensible
– Freely redistributable
– Open source

Some thoughts from yesterday
In the absence of infinite resources … need to think a bit harder.
• Joint efforts to reach critical mass, engagement from all.
• Coaching & training, at various levels – introductory, advanced.
• A new forum for discussion & exchange.
• The in-silico instrument scientist – some progress has been made.
• Databases & repositories – extension of INS database at ISIS?
• Get on with it - joint projects & demonstrators.
Demonstrate progress , and document it!

MDANSE 2016 SCHOOL TUTORIAL

Setting up a simulation with Forcite or DL_POLY

(Miguel A. Gonzalez, November 2016)

1. Introduction

There are many different software packages to perform classical Molecular Dynamics (MD) simulations, but here we will work either with the Forcite module in Materials Studio (Biovia) or DL_POLY. Each package has its own advantages and disadvantages. What follows is a personal view on the strengths and weaknesses of each code, but these opinions depend strongly on the different degree of knowledge (or lack of knowledge) of each code, so they should only be taken as indications for novice users. **If you are already familiar with another program or you collaborate with someone using a different MD package, my advice is that you keep using it as long as it fulfils your needs.**

Materials Studio and Forcite (http://accelrys.com/products/collaborative-science/biovia-materials-studio/)**:**

Materials Studio (MS) is a commercial package developed by Biovia and containing a large set of modules that allow you to do all kind of calculations (quantum mechanical methods, DFT, mesoscopic modelling, etc.). It provides an environment that makes very easy to set up a system from scratch, perform a MD simulation and then make some basic analysis.

This is clearly the choice for novice users desiring to try computer modelling without too much effort or to set-up and test quickly a simulation of a new system. The relevant tool in MS to perform classical MD simulations is Forcite, which is a collection of molecular mechanics tools allowing to perform energy calculations, geometry optimizations and molecular dynamics using a classical force field to represent the interactions between the nuclei.

Apart from the economic cost, the main drawback of MS is that it is not always easy (or even possible) to twist the program to perform less standard calculations. For example, if some parameters of the force field are missing or you want to test a particular potential from the literature that it is not included in the suite of potentials provided in Forcite, you will have to struggle to add it.

A more subtle problem is that it is extremely easy to setup a system and then simulate it, using the default options suggested by Forcite. While in many cases this will be OK, it is always wise to go through all the windows and check which are the options currently being used and try to understand what they mean and if they should be modified.

DL_POLY (http://www.ccp5.ac.uk/DL_POLY_CLASSIC/)**:**

This is a free parallel code that allows performing many different kind of simulations with a large set of potential functions, so the user has a large freedom to do what he really wants. It contains a GUI, but it is not particularly well developed, and in most cases the user will need to write the input files manually (although some utilities exist to help with this).

The input file containing the potential parameters is quite clear and the manual is very well written, so with some small effort a novice user should be able to write the needed input files to simulate a relatively small molecule. However preparing the input files for a complex system (e.g. a protein) may be quite hard.

2. Forcite

As said above, the tools provided by MS are combined in a very user friendly package, so hopefully creating a system of interest and start to perform some calculations on it will be relatively straightforward and you can start doing some interesting work just after playing a little bit with the program. Furthermore the package includes a series of well-prepared tutorials that you can follow to become familiar with the code. Some of the useful tutorials to learn how to setup a new system and perform a classical MD simulation on it are:

- Quick start tutorials: Create a project, sketch a molecule, and build a crystal or a polymer.
- Visualizer tutorials: More examples on sketching molecules/systems and using different builders.
- Amorphous Cell: This is the module needed to construct a 3D periodic box for an amorphous system.
- Forcite tutorials: Geometry optimization of urea (crystal).
- Forcite Plus tutorials: How to edit a force field, diffusivity of a gas in a polymer, miscibility of polymers.

As in the DL_POLY example we are going to build and simulate a box of water molecules, we can do the same using MS. To do this follow these steps:

- Open a new 3D Atomistic Document

- Select an oxygen atom in *Sketch Atom* (one of the many options in the icon menu bar) and then sketch the O atom in the document.

- Add the missing H atoms using *Adjust Hydrogen.*

- Open the *Amorphous Cell* module, select *Calculation* and in the energy tab select the COMPASS II force field. With this force field, the partial charges on the atoms are automatically assigned. Select the *Task: Construction*, give the number of water molecules that you want to put in the simulation box (e.g. 500) and the density (1 g/cm^3) and launch the calculation.

- Once that the calculation is finished, you will see a box of water molecules. You can check all the properties associated to the model that has been created (e.g. density, box size, box

contents, etc.) using the *Properties Explorer*. If you select an atom you can check the properties of the selected atom, in particular its *ForcefieldType* and its *Charge*. For example, you can check that the partial charge of an oxygen atom in the model created is –0.82 e, and that H atoms have a charge of +0.41 e.

- Now you can open the *Forcite* module and perform a first MD run. In the *Setup* tab choose *Task: Dynamics* and in the *Energy* tab *Forcefield: COMPASS II*. Then go back to the *Setup* menu and click on *More…* to open the dialog allowing to define the simulation conditions. We can start doing a NPT simulation to equilibrate the system at room conditions (298 K, 1 bar[1]). The time step can be let to its default value (1 fs) and we can set the total simulation time to 100 ps. Even if we are equilibrating the system and therefore the trajectory will not be used for analysis, we can save a snapshot of the instantaneous configuration of the system every 100 steps (i.e. 0.1 ps) to visualize later the trajectory and follow the equilibration of the system by computing any relevant property. For an NPT simulation we need to choose also a *Thermostat* and a *Barostat*. Now we can use for both the simple Berendsen thermostat and barostat with the default parameters. The *Job Control* tab allows to choose where to run the calculation (e.g. either in the computer where the MS interface runs or in a cluster, if available).

- Once the simulation is finished[2] we can look to the output file (named *Model.txt*) and check the final and average values of the different energy terms, the pressure, and the density. We can also see the time evolution of these quantities and check that the temperature reaches its nominal value after a few ps. The system volume also equilibrates in less than 10 ps and then the density fluctuates around the equilibrium value, so we have $\langle\rho\rangle \approx 0.97 \pm 0.01$ g/cm^3. You can also confirm that $\langle P\rangle \approx 0$ and that pressure fluctuations are very large, as we said ($\sigma(P) \approx 0.033$ GPa, i.e. ≈330 bar).

- Now we can use the final configuration to perform a production run in the NVT ensemble (to avoid later any inconvenience in the analysis[3]) and then use the *Analysis* tools available in the *Forcite* module to analyze the generated trajectory.

- As you can see in the next section, a simple but good model for water is the TIP4P/2005 potential, which uses different partial charges and Lennard-Jones parameters. We can apply the same charges to our model easily. Start by selecting all the oxygen atoms in the system using *Edit→Atom Selection→Select by Property: Element Is O* and then we can change the charge in the *Properties* explorer and set it to the desired value. Then we do the same for H.

- Setting a different pair of σ and ε values for the LJ potential is less straightforward. We can edit the potential using the *Forcefield Manager,* but not all available force fields can be edited. Thus we have to start by choosing a different force field. Open the document with the configuration of the final water box, launch the *Forcite* dialog and select *Forcefield:*

---

[1] You can specify 0.0001 GPa in the corresponding pressure box, but given the large fluctuations in pressure that we always have in the simulation (due to the small size of the simulated box compared to a macroscopic system one always have pressure fluctuations typically of a few hundred bars), letting P = 0 is equivalent.

[2] This simulation can take approximately 25 minutes using 16 cores.

[3] This is not compulsory, but some analysis programs may not work with simulation boxes that change during the trajectory.

Dreiding. Then perform simply an *Energy* calculation to associate the Dreiding force field types to the O and H atoms. Now if you select an O atom you will see that its *ForcefieldType* is O_3, while for H it is H___A. Then open the *Forcite Forcefield Manager* and select Dreiding among the Standard Forcefields. Go to the *Dreiding.off* window and click on *Filter by selection in:* and select the water model with the Dreiding forcefield types. We can modify now in the interface some of the default parameters of the Dreiding potential, such as bond stretching or angle bending force constants. However it is not possible to change directly the LJ parameters, so you need to open in an editor the Dreiding.off file, search for the force field type values under diagonal_VDW and modify them manually.

- Unfortunately the TIP4P/2005 model is a 4 site model in which the charge of the oxygen is not placed on the atom but in a 4th site of zero mass, and then the whole molecule is treated as a rigid body. This cannot be easily done in Forcite.

Continue playing by yourself to explore the different possibilities offered by MS and Forcite. You can follow some of the tutorials available or try to set up and simulate a system of interest to you.

3. DL_POLY

DL_POLY requires three files that must be compulsory named CONFIG, CONTROL, and FIELD.

- *CONFIG*: Contains the dimensions of the unit cell, a key to indicate the type of boundary conditions applied, and the coordinates of each atom (and eventually also velocities and forces).
- *FIELD*: Contains all the force field information defining the system to be simulated and the nature of the interatomic molecular forces.
- *CONTROL*: Contains all the directives controlling the simulation.

The program does not really have any UI allowing to prepare easily those files, so you will need to manage by yourself to create the three files following the needed specifications. In the program distribution there is a Java interface that could be used as a starting point to prepare such files and also contains some analysis tools to analyze the trajectories generated by DL_POLY, but the possibilities offered by this interface are somewhat limited. A very interesting alternative seems to be the Aten project (https://www.projectaten.com/) by Tristan Youngs (ISIS, STFC, UK).

In the example here we are going to simulate water, so the system is simple enough to allow us to create the three input files manually. However we are going to use another utility called DL_FIELD[4] to help us with this, as this program can be very useful when trying to set up more complex systems.

[4] http://www.scd.stfc.ac.uk//research/app/ccg/software/DL_FIELD/40633.aspx. DL_FIELD is free for academic scientists, but requires registration, so if you are interested in using it, please register by following the instructions in this webpage and get the latest version. Here we have used DL_FIELD 3.1.

Before starting any simulation we need to have a clear idea of which is the model that we want to study. For example, here we are interested in water. However we have to keep in mind that in our simulation we are not going to study real water, but a particular model of water. As water is probably the most widely studied system both experimentally and computationally, there are many available models in the literature (see e.g. http://www1.lsbu.ac.uk/water/water_models.html). Here we will explore a relatively recent model proposed as an improvement of the older TIP4P potential [1]. This TIP4P/2005 potential is a very simple model that represents H_2O as a **rigid** molecule having four sites (the three atoms + an additional site to place the negative charge associated to the lone pair (LP) of the oxygen). The interaction between a pair of water molecules is given by the sum of the electrostatic interactions between the partial charges placed in the H atoms and the LP site plus a Lennard-Jones potential acting between the O sites. The geometry of the molecule and the model parameters are:

d(O-H) = 0.9572 Å, θ(H-O-H) = 104.52°, d(O-LP) = 0.1546 Å;

q(LP) = –1.1128 e, q(H) = +0.5564 e

σ = 3.1589 Å, ε = 93.2 k_B

In spite of its simplicity, the TIP4P/2005 potential has proved very successful in reproducing many properties of water [2], so it seems a reasonable choice to simulate liquid water.

The only file that we need to start is a PDB file. PDB stands for Protein Data Bank and is a file format that provides a convenient representation for macromolecular structures. As the name indicates it has been originally designed to represent protein configurations, but it has become one of the most used format for all kind of systems. One can easily find a PDB file of a water configuration in the net, but naturally most of them will contain the positions only of the O and H atoms, so the LP pair will be missing. Furthermore the model that we are going to use is a rigid one, so all the water molecules in our box should have exactly the same geometry given by the bond lengths and angles above. Therefore in this case it is more convenient to calculate the coordinates of the 4 sites and write manually the PDB file[5], which could look like this:

```
ATOM    1  O1  TIP4   1      0.000   0.000   0.000  1.00  0.00   tip4p   O
ATOM    2  H2  TIP4   1      0.586   0.757   0.000  1.00  0.00   tip4p   H
ATOM    3  H3  TIP4   1      0.586  -0.757   0.000  1.00  0.00   tip4p   H
ATOM    4  Q4  TIP4   1      0.155   0.000   0.000  1.00  0.00   tip4p   Q4
END
```

Save this as e.g. *water_1.pdb*. Now we need to check the water models available in the DL_FIELD library. For example open the file *DLPOLY_CHARMM.sf* (in *dl_field_3.1\lib*) and search for *MOLECULE_TYPE*. You can see there the list of molecules that have been defined in the file. In the case that your molecule is not in the list, then you will need to add it using a similar molecule as reference either in the same file or in a user-defined force field (see the DL_FIELD manual for more details). As you can see there are several water models available, including *water_tip4p*. This is not the TIP4P/2005 model that we want to use, but

[5] A detailed explanation of the PDB format can be found in **http://www.rcsb.org/pdb/static.do?p=file_formats/pdb/index.html**

the original one having slightly different charges, LJ parameters on the O and intramolecular distances. We could add a new type to the forcefield file, but for the moment we will just use the water_tip4p as given in *DLPOLY_CHARMM.sf* and modify later the resulting FIELD file.

```
MOLECULE water_tip4p 4 0.0      TIP4P with pseudo atom Q4
O    O_tip4p       0.00
H1   H_tip4p       0.52
H2   H_tip4p       0.52
Q4   pseudo_tip4p -1.04
CONNECT O   > H1 H2
CONNECT H1  > O
CONNECT H2  > O
CONNECT Q4  > Q4
RIGID O H1 H2 Q4
END MOLECULE TIP4P
```

Now we have to edit the dl_field.control containing the directives for DL_FIELD. We need to indicate which of the FF available in the library we want to use (e.g. charmm in this example), the name of the PDB file containing the positions of our system (water_1.pdb) and in this case that we have a rigid body (option Apply rigid body? set to 1). When this option is used, then we need to add also the RIGID directive and the name of the molecular group that will be treated as a rigid body. This name is the Molecular group name which in the PDB file appears in columns 70-76.

```
This is the title line. Reads only 80 columns
charmm    * Type of force field require (see list below for choices).
kJ/mol    * Energy unit: kcal/mol, kJ/mol, eV or default.
0         * Bond type (0=default, 1=harmonic , 2=Morse)
0         * Angle type (0=default, 1=harmonic, 2=harmonic cos)
0         * Include user-defined information (dl_field.udff): 1=yes 0=no
1         * Verbosity mode: 1 = on, 0 = off
water_1.pdb       * Configuration file.
none      * Output file in PDB. Put 'none' if not needed.
1         * Optimise FIELD output size, if possible? 1=yes  0=no
2         * Atom display: 1 = DL_FIELD format. 2 = Standard format
2         * Vdw display format: 1 = 12-6 format   2 = LJ format
1         * Display additional info. for protein 1=Yes  0=No
0         * Freeze atoms? 1 = Yes (see below)  0 = No
0         * Tether atoms? 1 = Yes (see below)  0 = No
0         * Constrain bonds? 1 = Yes (see below) 0 = No
1         * Apply rigid body? 1 = Yes (see below) 0 = No
1         * Periodic condition ? 0=no, other number = type of box (see below)
4.00  0.00  0.00  * Cell vector a (x, y, z)
0.00  4.00  0.00  * Cell vector b (x, y, z)
0.00  0.00  4.00  * Cell vector c (x, y, z)
default   * 1-4 scaling for coulombic (put default or x for scaling=x)
default   * 1-4 scaling for vdw (put default or x for scaling=x)
0  300.0  * Include velocity? 1=yes, 0=no and scaling temperature.
1         * Position solute at origin? 1 = yes, 0=no
none  1.6 * Solvate model? none or specify solvent (see below) and distance criteria.
0  20.0   * Add counter ions? 1=yes, 0=no,  minimum distance from solute
0         * MM energy calculation. 1=Yes, 0=No
```

```
12.0      * Cut off for electrostatic energy calculation (angstrom)
12.0      * Cut off for vdw energy calculation (angstrom)

#########################################################
RIGID     tip4p
#########################################################
```

Running now dl_field_3.1 will write the three files *dl_field.CONFIG, dl_field.CONTROL* and *dl_field.FIELD* in the output directory. We only have one water molecule in our system, but we can use the *mfold.f* program in dl_class_1.9/utility to replicate the system 8 times in each direction in order to create a larger simulation box of side L = 32 Å and containing 512 water molecules. Rename this file to CONFIG.

Now we have to edit dl_field.FIELD and change some values. First change *nummols* from 1 to 512 as now we have created a larger box containing 512 water molecules. Then change the charges on the two H atoms from 0.54 to 0.5564 and the charge in the *Q4* site from –1.04 to –1.1128. Finally modify the LJ parameters for the OO interaction to be ε=0.7749 kJ/mol and σ=3.1589 Å. Your file should look similar to this:

```
Generated by DL_FIELD v3.10
Units kJ/mol
Molecular types 1
Molecule name tip4p
nummols 512
atoms 4
OT4        15.99940     0.00000   1   0    1TIP4
HT4         1.00797     0.55640   1   0    1TIP4
HT4         1.00797     0.55640   1   0    1TIP4
Q4         0.00000    -1.11280   1   0    1TIP4
rigid 1
4  1 2 3 4
finish
vdw 6
OT4     OT4     lj       0.7749      3.1589
HT4     OT4     lj      -0.0000      1.5768
Q4      OT4     lj      -0.0000      1.5768
HT4     HT4     lj       0.0000      0.0000
Q4      HT4     lj       0.0000      0.0000
Q4      Q4      lj       0.0000      0.0000
Close
```

Save it as FIELD and create a new CONTROL file containing the following commands:

```
TIP4P/2005 water NPT, T=298 K
integrator leapfrog verlet
temperature 298.0 K
pressure    0.001 kbar
ensemble npt ber 1.0 1.0
steps 10000
print 10
stack 10
stats 10
timestep  0.002
cutoff   12.0
```

```
delr      2.0
spme precision  1d-6
job time 500000.0
close time 100.0
finish
```

This will perform a short (20 ps) NPT simulation to start equilibrating the system. As we have started very far from equilibrium conditions (very low initial density and with an ordered system as it has been created by replicating a single molecule) this will not be enough to really equilibrate the system, but performing longer simulations should bring our model progressively to equilibrium.

As we have our 3 files, we can run dl_class_1.9/execute/DLPOLY.X and check that everything works. Using a single processor this run can last about 10 minutes and the box size should decrease from 32 to ≈27 Å. This corresponds to a molecule number density ≈0.026, indicating us that we need to continue our run to approach the expected value of 0.033 molecules/Å^3.

You can continue the simulation of this system until it is well equilibrated[6] and then perform a production run in the NVT or the NVE ensemble. While doing this you can continue working with DL_FIELD to prepare other systems.

4. References

[1] J.L.F. Abascal and C. Vega, J. Chem. Phys. 123, 234505 (2005)

[2] C. Vega and J.L.F. Abascal, Phys. Chem. Chem. Phys. 13, 19663 (2011)

[6] At this point, the volume, temperature, pressure, and potential energy should fluctuate around their average values.

MDANSE 2016 SCHOOL TUTORIAL

Analysis of the structure and dynamics of water using MDANSE

(Eric Pellegrini and Miguel A. Gonzalez, November 2016)

1. Introduction

We are going to use MDANSE[1] to show how to analyze a typical Molecular Dynamics trajectory corresponding to liquid water at ambient conditions. Water is probably the most widely studied system both experimentally and computationally. But nevertheless a large number of questions related to their anomalous properties remain. From a simulation perspective, many different models have been proposed in order to reproduce the properties of liquid water and ice from MD simulations. Here we will explore a model proposed recently as an improvement of the older TIP4P potential [1][2]. This **TIP4P/2005** potential is a very simple model that represents H_2O as a **rigid** molecule having four sites (the three atoms + an additional site to place the negative charge associated to the lone pair (LP) of the oxygen). The interaction between a pair of water molecules is given by the sum of the electrostatic interactions between the partial charges placed in the H atoms and the LP site plus a Lennard-Jones potential acting between the O sites. The geometry of the molecule and the model parameters are:

d(O-H) = 0.9572 Å, θ(H-O-H) = 104.52°, d(O-LP) = 0.1546 Å;

q(LP) = –1.1128 e, q(H) = +0.5564 e

σ = 3.1589 Å, ε = 93.2 k_B

In spite of its simplicity, this model has proved very successful in reproducing many properties of water [2]. Here it has been employed to simulate a system containing 360 water molecules at ambient conditions.

The system has been simulated for 100 ps using DL_POLY Classic 1.9 and the instantaneous positions saved every 0.2 ps. The input files used (FIELD, CONFIG, CONTROL) and the output files generated by the simulation (OUTPUT, HISTORY, STATIS, REVCON, REVIVE) are provided in the tutorial folder. The average thermodynamic results calculated by the program are given at the end of the OUTPUT file (open it with an editor and search 'final averages'). Important parameters that should be given when describing the results of the simulation in a publication are:

$\langle T \rangle = 298 \pm 9$ K

n = 360/V = 0.03304 molecules/Å^3 → d = 0.988 g/cm^3

$\langle P \rangle = -0.2 \pm 0.7$ kbar

[1] This in an interactive application for analysing molecular dynamics simulations. It is still under development, but a beta version (currently 1.0.4.rc2 on Nov 30th 2016) is already available from https://mdanse.org.
[2] References are provided in the tutorial folder as Ref1.pdf, Ref2.pdf, etc.

$\langle U \rangle$ = –17170 ± 143 kJ/mol (the output of DL_POLY corresponds to 1 mol of simulation boxes, i.e. –47.7 ± 0.4 kJ/mol of water molecules)

2. Convert the trajectory

As MDANSE can only work with MMTK trajectories, we first need to convert the trajectory generated by our MD favourite program from its original format to the MMTK trajectory format. MDANSE can convert trajectories generated both by DFT codes such as CASTEP, VASP, DMOL3 and DFTB and by classical MD codes, such as CHARMM, X-PLOR, NAMD, DL_POLY, LAMMPS, Discover, and Forcite. In case a different code has been used to generate the trajectory, it is possible to use either the PDB or the Generic converters once the user has converted the initial trajectory to one of those simple formats.

In our example we need to open the trajectory converter for DL_POLY and we see that it requires two files: FIELD (this is an input file for DL_POLY containing the system definition) and HISTORY (the output file containing the trajectory generated by DL_POLY). Additionally the interface shows a box named 'Atom aliases' that allows the user to rename one or several atoms whenever this is needed. The reason is that the converter needs to use the atom name given in the FIELD file to guess which element it is. In most cases, this will work fine and names such as CB, CC, HA, HB, HC will be correctly interpreted as carbon and hydrogen. But if we have simulated a system containing an OH group and named the hydroxyl hydrogen as HO, or we have several carbons named CA, the converter will interpret them as holmium and calcium, respectively. In order to avoid this, in these cases the user needs to provide the needed information as a python dictionary in the 'Atom aliases' box, e.g.: {'HO':'H', 'CA':'C'}. If we check the atom names in FIELD we see that they are named 'O' and 'H', so they will be correctly interpreted. The additional site corresponding to the lone pair of the oxygen is named 'Du', which will also be correctly interpreted as there are already a dummy (symbol 'Du') and a center (symbol 'Cen') atoms in the elements database of MDANSE.

Once the trajectory has been converted, it can be loaded in MDANSE and the plugins box will be updated to show the available operations. It is often useful to check the content of the file in *' Miscellaneous: Data Info'* to get some basic information and verify that the conversion was succesful and use *'Viewer: Molecular Viewer'* to inspect visually the trajectory (Fig. 1).

Now we will explore some of the analysis available in MDANSE.

3. Analysis: Thermodynamics

The two analysis here give the instantaneous density and temperature of the simulated system. In most cases, this information is already given directly by the MD software, but otherwise MDANSE can be used to explore its evolution during the simulation.

Run both analysis and use the NetCDF plotter (4th symbol in the graphical menu bar) to visualize the result. The mass and atomic density are constant all along the simulation (Fig. 2). **Why?**

The result for the mass density is 0.98831 g/cm^3, as calculated above from the volume of the simulated box. However we also calculated before that the molecular density is 0.03304 molecules/Å^3, i.e. 0.09912 atoms/Å^3, but MDANSE gives 1.3215×10^{23} atoms/cm^3. **Why?**

Now check the temperature. The instantaneous temperature fluctuates, as expected, but its value is far off from the expected value of 298 K corresponding to the average temperature of our simulation (Fig. 4). **Why?** Hint: Try using different values for the interpolation order. This will not really improve the result, but it can provide a hint to the reason why the analysis fails and to what would be needed to get the correct temperature.

4. Analysis: Structure

One of the important steps to validate our simulation is to check that the model employed is able to reproduce reasonably well the structure of the real system. Often this is done by comparing the results of the simulation with experimental data obtained by means of neutron or x-ray diffraction.

A first useful analysis to get an insight into the structure of our model can be obtained using the analysis *'Pair Distribution Function' (PDF)*. Launch the analysis and give the following *R values: from 0 to 1.1 by step of 0.01*. All the input values must be given in the natural units used by MDANSE (ps for time, ps^{-1} for angular frequency, nm for distance, and nm^{-1} for *Q*), so this means that we will calculate the pair distribution function up to a maximum distance of 1.1 nm and using a step of 0.01 nm in our histogram. You should also note that the maximum distance should not be larger than the radius of the largest sphere that can be circumscribed by the simulation box. MDANSE will accept larger values and perform the calculation up to the value given, but the result will not be correct beyond that distance. **Why?** Hint: Run the analysis using a value larger than 1.11 (here we used a cubic simulation box of side 22.17 Å) and look to the result (Fig. 6).

Almost all the analysis in MDANSE allow the user to select a subset of the full system using the *'Atom selection'* box. Additionally the box *'Atom transmutation'* permits to change the selected atoms into another isotope or even another atom. For example, in the case of water most of the neutron diffraction experiments have been done using heavy water to avoid the large incoherent scattering from hydrogen. Therefore in order to compare with experiment we may select all the hydrogen atoms and transmute them into 'h2'. The final quantity that we must select is the weight given to each atom to compute its contribution to the total property. For most analysis, the value by default here is 'equal', implying that all atoms contribute equally to the total. However the user can change this and select any of the properties available in the database. And as we will see later, for some analysis the default is set to a different value, as b_{coh} for the Static Structure Factor or the Dynamic

Coherent Structure Factor analysis or b_{inc}^2 for the Dynamic Incoherent Structure Factor analysis.

After running the PDF analysis, you can open the output file in the 2D/3D Plotter and check its contents. You will see that both the intramolecular and intermolecular PDFs are computed for each possible pair of atoms (including the dummy atom), as well as their sum (total) and the sum over all the pairs. The latter is computed using the following expression:

$$w_{\alpha\beta} = \frac{N_\alpha N_\beta f_\alpha f_\beta}{\sum_i N_i N_j |f_i||f_j|}$$

where $N_{\alpha(\beta)}$ are the number of atoms of type $\alpha(\beta)$ and $f_{\alpha(\beta)}$ are the weighting factors mentioned above.

You can plot the relevant PDFs (OO, OH, and HH) and export them into text files to manipulate or visualize them outside MDANSE.

Tables of site-site radial distribution functions for water at 300K obtained from experimental measurements (both neutron and x-ray diffraction) combined with an EPSR (Empirical Potential Structure Refinement) analysis have been published by A. K. Soper [3]. The values of the radial distribution function of Table 5 of this paper are in the file *soper13_water_structure_review_rdf.dat*. You can compare them with the 3 intermolecular PDFs given by MDANSE[3] (Figs. 7-9).

Another quantity of interest is the coordination number. You can compute it and check the value of the oxygen-oxygen coordination number around the first minimum of the O-O p.d.f. (Fig. 10). What does it tell us about the structure of water?

Now use the *Static Structure Factor* analysis to compute the S(Q), which is the quantity that will be directly measured in a diffraction experiment. In the case of MDANSE, it is calculated as the Fourier Transform of the p.d.f., so the interface requires the same *R values* as for the PDF analysis and in addition the Q-range and Q-step over which to compute S(Q). Again there are no limits to the minimum and maximum Q values, but you should note that the value of S(Q) will be meaningless for $Q < 2\pi/R_{max}$, where is the maximum R given above. Thus a reasonable choice here could be *Q values: from 6 to 200 by step of 1* (remember that the units here will be nm^{-1}). You will notice that for this analysis the default weight is *'b_coherent'*, so you need to change the hydrogen atoms to 'h2' in order to be able to compare directly the output of the calculation with the results obtained from diffraction measurements on D_2O.

Concerning the partial static structure factors, one should note that there are different definitions that can be employed to define them, notably those given by Faber-Ziman and Ashcroft-Langreth (e.g. see equations (2.35) and (2.36) in [4]). MDANSE uses the Faber-Ziman definition:

[3] If you are familiar with Python and Matplotlib (or you simply want to try it) you can use the files fig_pdf_oo(hh)(oh).py to do this.

$$S_{\alpha\beta} = 1 + \frac{4\pi\rho_0}{Q}\int_0^\infty r\left[g_{\alpha\beta}(r) - 1\right]\sin(Qr)\,dr$$

where ρ_0 is the total number density of atoms.

The total (intra- + inter-molecular) and the intermolecular partial static structure factors can be compared to the experimental data shown in Figs. 6-8 of [5].

The total static structure factor is calculated using the standard weighting scheme presented above. As in this analysis we have used the coherent scattering lengths as weights, we have:

$$S(Q) = \frac{\sum_{\alpha\beta}^{n} c_\alpha c_\beta b_\alpha b_\beta\, S_{\alpha\beta}(Q)}{\sum_{\alpha\beta}^{n} c_\alpha c_\beta |b_\alpha||b_\beta|}$$

In order to compare this result to experiment we may refer to ref. [3], where A. K. Soper gives experimental data for the differential scattering cross section, $D_n(Q)$, of D_2O in table 3[4]. The definition of $D_n(Q)$ is:

$$D_n(Q) = I_n(Q) - \sum_{\alpha}^{n} c_\alpha \langle b_\alpha^2\rangle = \left(\sum_{\alpha\beta}^{n} c_\alpha c_\beta b_\alpha b_\beta\right)[S(Q) - 1]$$

and using $\left(\sum_{\alpha\beta}^{n} c_\alpha c_\beta b_\alpha b_\beta\right) = 0.4076\ \mathrm{barn/sr/atom}$ for heavy water we can compare the simulated and measured static structure factors.

We can also compute the static structure factor that will be obtained using x-rays instead of neutrons. One possibility is to repeat the same analysis but using the atomic numbers instead of b_{coh} as weighting factors. However this will neglect the Q-dependence of the atomic form factors. Thus a more accurate solution is to use the *XRay Static Structure Factor* analysis, which computes the partial $S_{\alpha\beta}(Q)$'s and calculates the total one using the appropriate atomic form factors.

After doing this, we also need to pay attention to the normalization procedure used in the treatment of the experimental data (see again Ref. [3]). For example, for normalization II in [3] we have:

$$D_x(Q) = \frac{I_x(Q)}{\sum_\alpha c_\alpha f_\alpha^2(Q)} - 1$$

with

$$I_x(Q) = \sum_{\alpha} c_\alpha f_\alpha^2(Q) + \sum_{\alpha\beta\geq\alpha}\left(2 - \delta_{\alpha\beta}\right) \times c_\alpha c_\beta f_\alpha(Q) f_\beta(Q) H_{\alpha\beta}(Q)$$

$$H_{\alpha\beta}(Q) = 4\pi\rho\int_0^\infty r^2\left[g_{\alpha\beta}(r) - 1\right]\frac{\sin(Qr)}{Qr}dr = S_{\alpha\beta}(Q) - 1$$

[4] These data are given in file *soper13_water_structure_review_dq_neutron.dat.*

giving:

$$D_x(Q) = [S(Q) - 1] \times \frac{\sum_{\alpha\beta} c_\alpha c_\beta f_\alpha(Q) f_\beta(Q)}{\sum_\alpha c_\alpha f_\alpha^2(Q)}$$

where S(Q) is the total X-ray static structure factor given by MDANSE. The experimental data given in Table 4 of [3] are given in the file *FxQ_water_Soper.dat* and the python script *fig_ssf_h2o_xray.py* gives an example of how to compare experiment and simulation using the equation above (Fig. 11).

5. Analysis: Dynamics

Another of the standard quantities typically computed for each simulation is the Mean Square Displacement (MSD). The slope of the m.s.d. at long times is proportional to the self-diffusion coefficient: $u^2(t \to \infty) \propto 6Dt$, so compute the MSD and then analyze its output to determine *D* and compare it to the experimental value for water at room temperature (2.3 $\times 10^{-9}$ m^2/s)[5] (Fig. 12).

In some cases it can be useful to compute the MSD (or any other property) for the centre of mass of the molecule. It is possible to create a new trajectory containing only the positions of the centre of masses of each molecule using the tool *Trajectory: Center Of Masses Trajectory* and selecting *'Group coordinates by molecule'*. Now we can load this new trajectory and compute the MSD for the molecular centre of masses, although in this case the result is naturally extremely close to the MSD of the oxygen atoms.

The velocity autocorrelation function and its Fourier transform (the density of states) are also of theoretical interest, but in order to be able to compute them with the necessary accuracy we need a trajectory containing also the instantaneous velocities or where the time interval between two consecutive frames is short enough to allow a reasonable determination of those.

6. Analysis: Scattering

For a neutron scattering experimentalist, two of the more useful analysis will be the *Dynamic Incoherent Structure Factor* and the *Dynamic Coherent Structure Factor.*

In both cases we need to define a set of Q vectors for which the F(Q,t) and S(Q,ω) will be calculated. The *Q vectors* interface allows to select between different possibilities depending on the type of system that we are investigating (e.g. many of the available options are only needed to study specific directions in a crystal), but for an isotropic system we can use the *spherical* option. It should also be noted that for the incoherent case there is not any restriction to the possible vectors, but in the coherent one the only acceptable vectors must be lattice vectors, so we need to select the option *spherical_lattice*. Use this

[5] You can use the file fig_msd.py

option and generate a set of vectors between 0.4 and 2 $Å^{-1}$ by giving *Shells from 4 to 20 by step of 4* (nm^{-1}). In order to accelerate the computation we can decrease the number of *N vectors* (e.g. to 10). Finally the *width* field allows to change the tolerance to accept vectors whose modules are more or less close to the nominal value, e.g. using *width = 1* we will generate vectors in the ranges (0.35-0.45), (0.75-0.85), (1.15-1.25), (1.55-1.65), and (1.95-2.05) $Å^{-1}$.

The other parameter that we need to set is the instrument resolution, which plays a double role. First, it is used to generate a smoothing function that will multiply the intermediate function computed directly in the analysis before doing the Fourier transform in order to avoid numerical artifacts. Second, if we have experimental data measured in a particular instrument we can choose it to correspond to the resolution of this instrument in order to facilitate the comparison between the simulated and measured spectra. Several resolution functions can be chosen, but the most common one is the Gaussian, for which the position (normally 0) and the value of σ must be given. The latter is given in units of ps^{-1}, which can be converted to meV by multiplying by 0.6582. As the full-width at half-maximum (FWHM) of a Gaussian function is 2.35842σ, we have $\sigma \approx$ FWHM (in meV)/1.55. For example, in order to compare with data taken in IN5 (a Time-of-Flight spectrometer at the ILL) using λ_0 = 5 Å (FWHM ≈ 0.09 meV) (see [7]) we may use σ = 0.06.

The file *IN5_H2O.hdf* contains some experimental data for pure water measured close to ambient conditions (T = 293-295K) in IN5 using λ_0 = 5 Å. The measured data have been corrected following standard procedures and interpolated into a constant Q-ω grid in the range $0.3 \leq Q \leq 2.1$ $Å^{-1}$. We can try now to compute the incoherent $S(Q,\omega)$ in the same Q-range and compare simulation to experiment (Fig. 13). The self-intermediate function F(Q,t) can also be compared to the results shown in [7], which is a very good example of the information that it is possible to extract from the simulation.

Now we can follow the same procedure to compute the coherent $S(Q,\omega)$ which could be compared to data measured in IN5 for D_2O. But as you will see the comparison is not as straightforward, as the simulated dynamic structure factor is very noisy. The reason is that when computing the incoherent dynamic structure factor we are calculating it for each individual atom and then averaging over N atoms, while we cannot do any averaging when calculating the coherent one. Therefore the statistical quality of the correlation functions is approximately sqrt(N) times better for $F_{inc}(Q,t)$ than for $F_{coh}(Q,t)$. The *Water_trajectories* folder contains other 5 equivalent trajectories. You can analyze and compare them to get an idea of the magnitude of the errors done when computing different properties (e.g. PDF, MSD, DISF, DCSF) and as a function of time (e.g. compare how curves for MSD, DISF, DCSF diverge) (Figs. 14-19).

The *Experiment* folder contains some experimental data measured in IN5 for H_2O, D_2O, CH_3OH, CH_3OD, and CD_3OH, while in *Methanol_MDANSE* there is a methanol trajectory. You can try to use MDANSE to simulate the differences between different isotopic substitutions and compare to the experimental data.

7. References

[1] J.L.F. Abascal and C. Vega, *J. Chem. Phys.* **123**, 234505 (2005).

[2] C. Vega and J.L.F. Abascal, *Phys. Chem. Chem. Phys.* **13**, 19663 (2011).

[3] A. K. Soper, ISRN Physical Chemistry Volume 2013, Article ID 279463 (2013).

[4] H. E. Fischer, A. C. Barnes, and P. S. Salmon, *Rep. Prog. Phys.* **69**, 233 (2006).

[5] A. K. Soper and M. G. Phillips, *Chem. Phys.* **107**, 47 (1986).

[6] R. Mills, *J. Phys. Chem.* **77**, 685 (1973).

[7] J. Qvist, H. Schober, and B. Halle, *J. Chem. Phys.* **134**, 144508(2011).

Appendix: Figures and answers

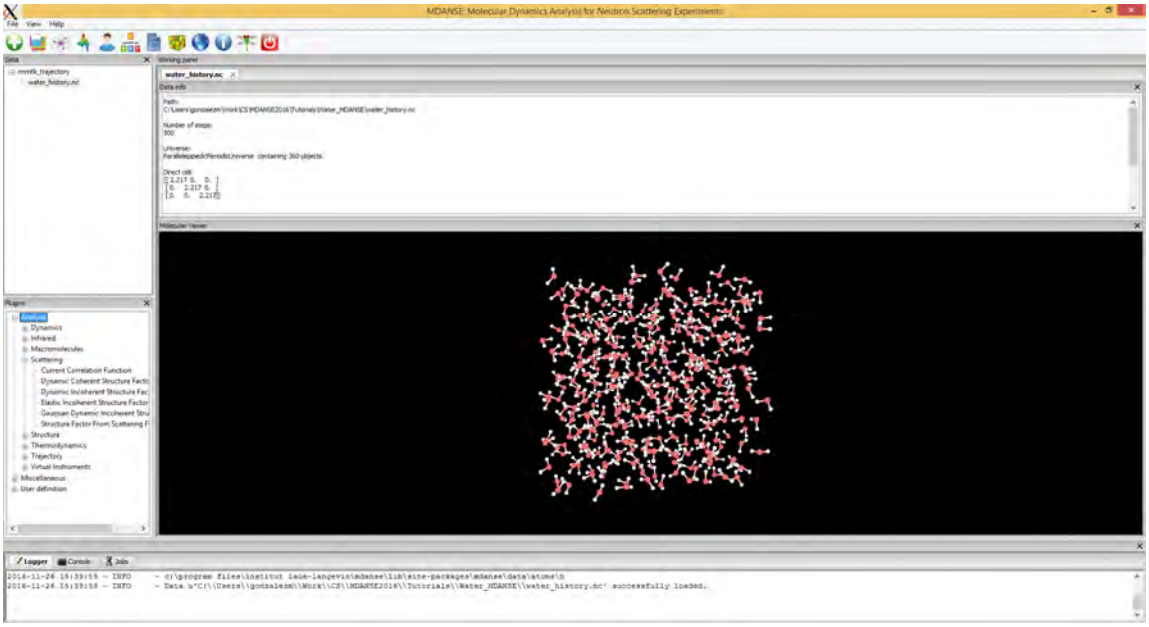


Fig. 1: Main interface of MDANSE showing the molecular viewer, the data info window and the available analysis.

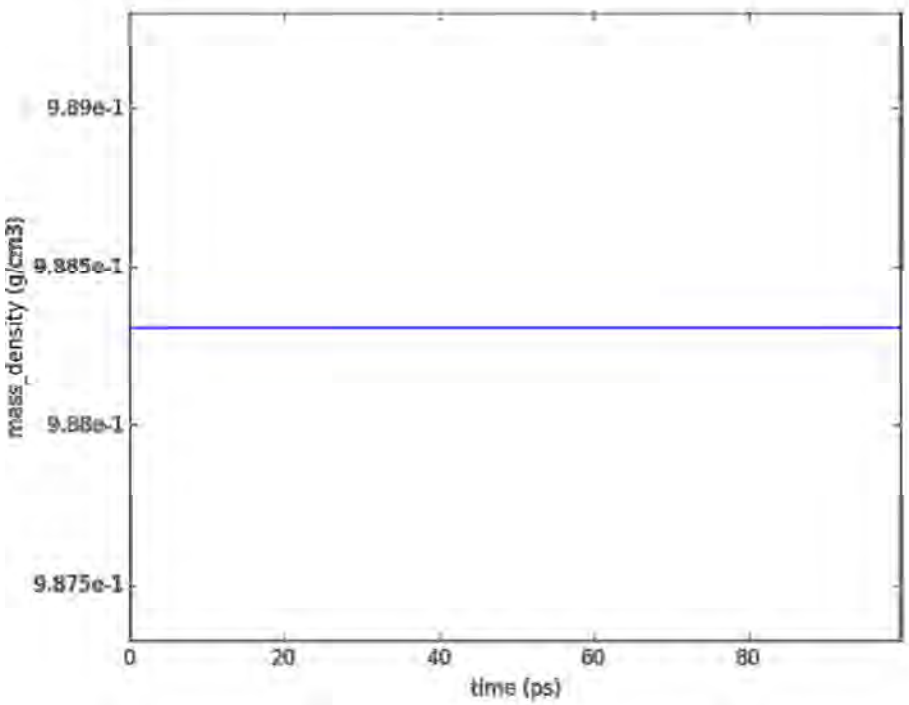


Fig. 2: Mass density of the simulated water box. As we have performed a NVT simulation, both the number of molecules and the volume of the simulation box are constant, so ρ is also a constant. The atomic density is calculated using the total number of sites of the model, so in this case the lone pair site is incorrectly taken into account in the calculation. A similar problem will happen if our model contains united atoms (i.e. a site representing more than one atom, e.g. a CH_3 or a CH_2 group).

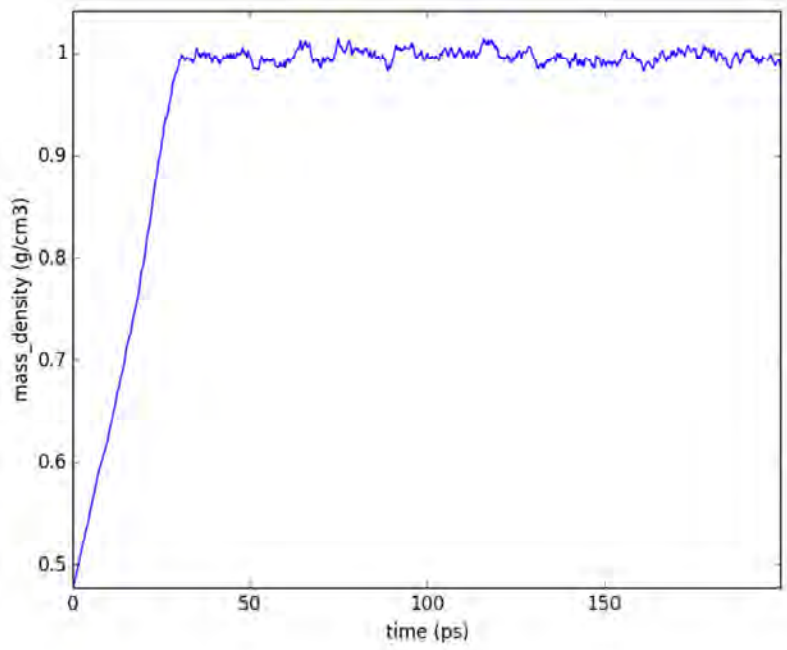


Fig. 3: Example of the evolution of the mass density during a NPT simulation of liquid water. In this case the starting configuration consisted in a box of 512 water molecules built by replicating a box of 4 Å side 8 times in each direction. The initial system is clearly not a good representation of liquid water and the initial density is very low. However after ≈30 ps of simulation in the NPT ensemble (using Berendsen's thermostat and barostat with a coupling constant of 1 ps) the system has reached the expected density for H_2O at ambient conditions.

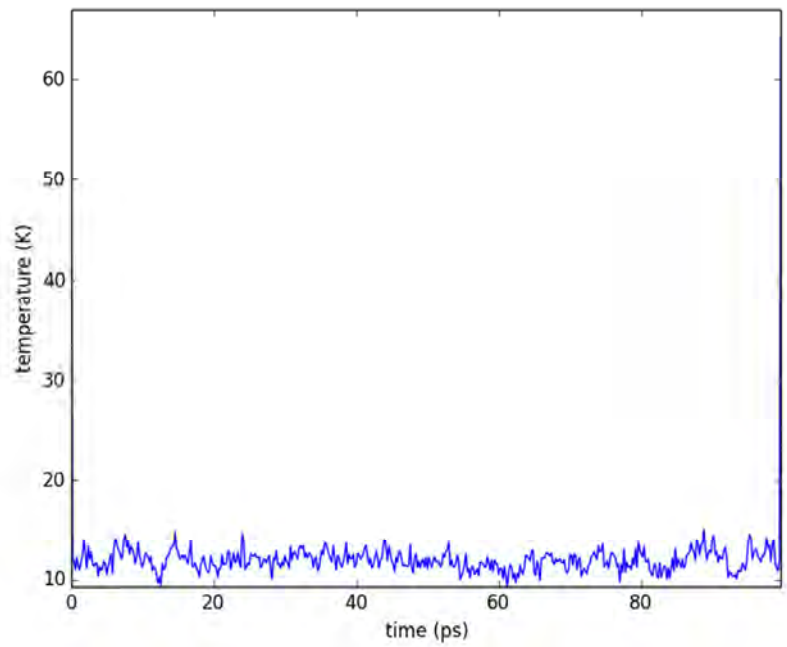


Fig. 4: Instantaneous temperature. The result is far off from the expected value and the average temperature of the simulation given in the OUTPUT file. The reason for this is that T is computed from the kinetic energy using the equipartition theorem, but as the velocities are not present in the trajectory file they have to be computed from the derivative of the positions. As the time interval between two consecutive saved configurations is not sufficiently short to compute the velocities with enough accuracy, the temperature cannot be correctly estimated.

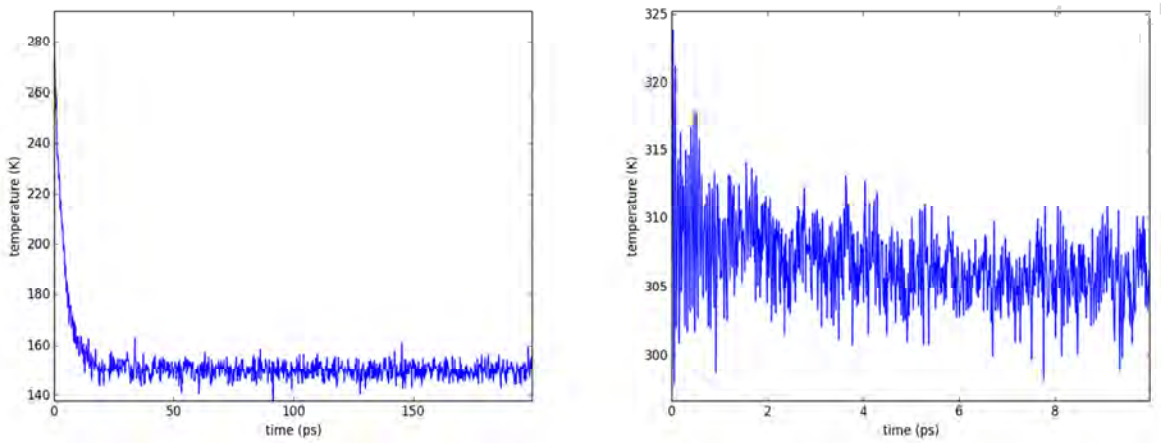


Fig. 5: (left) Temperature evolution during the NPT equilibration run for water of Fig. 3. Now the instantaneous velocities were also saved, so the calculation of T is more precise. However we do not get yet the correct result (≈ 298 K) because the algorithm is not directly applicable to rigid models. (right) Example of the same calculation on a fully flexible model of methanol.

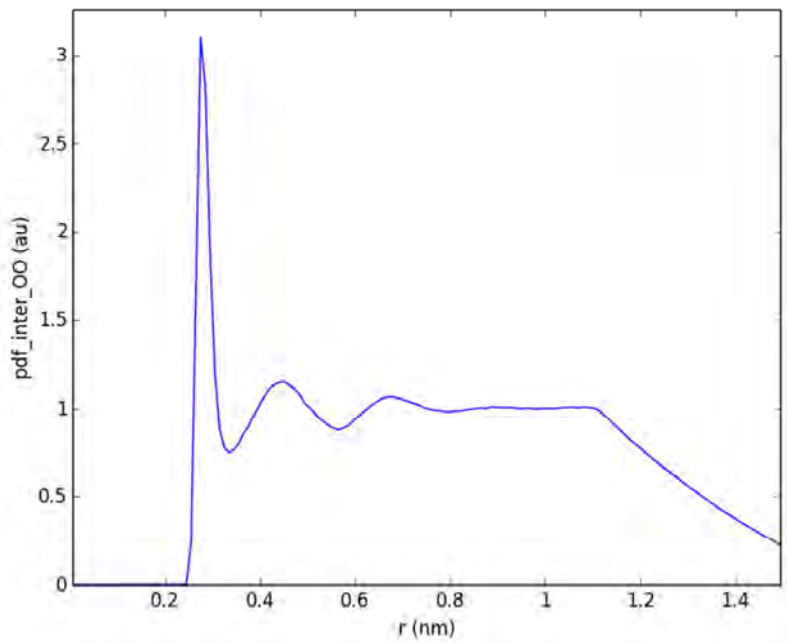


Fig. 6: Oxygen-oxygen pair distribution function. As only the closest image is used in the analysis, beyond L/2 the spherical shells over which the local density is calculated are 'incomplete' and the pdf decreases unphysically.

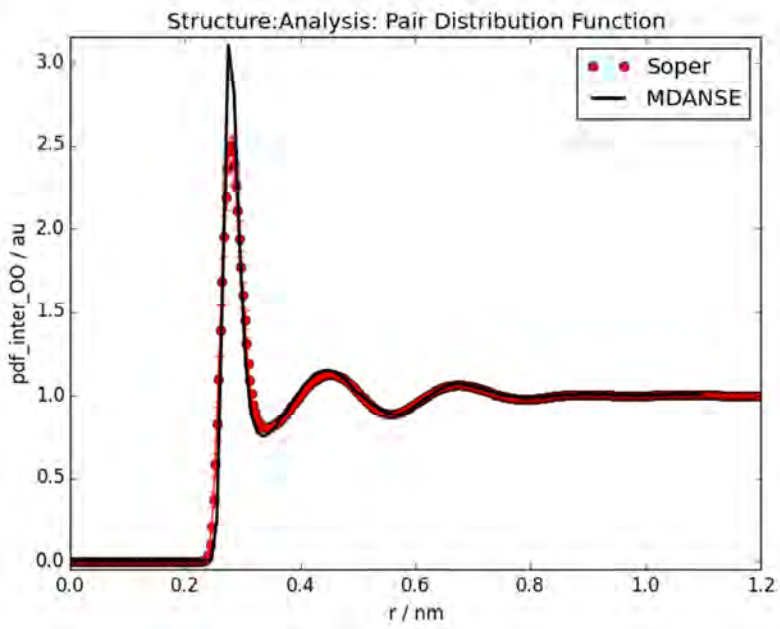


Fig. 7: Oxygen-oxygen pair distribution function: Simulation vs experiment.

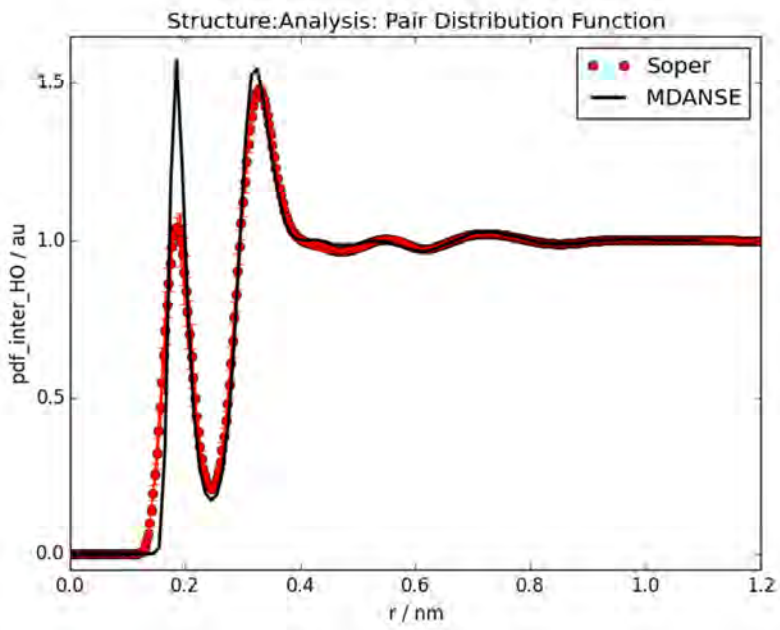


Fig. 8: Oxygen-hydrogen pair distribution function: Simulation vs experiment.

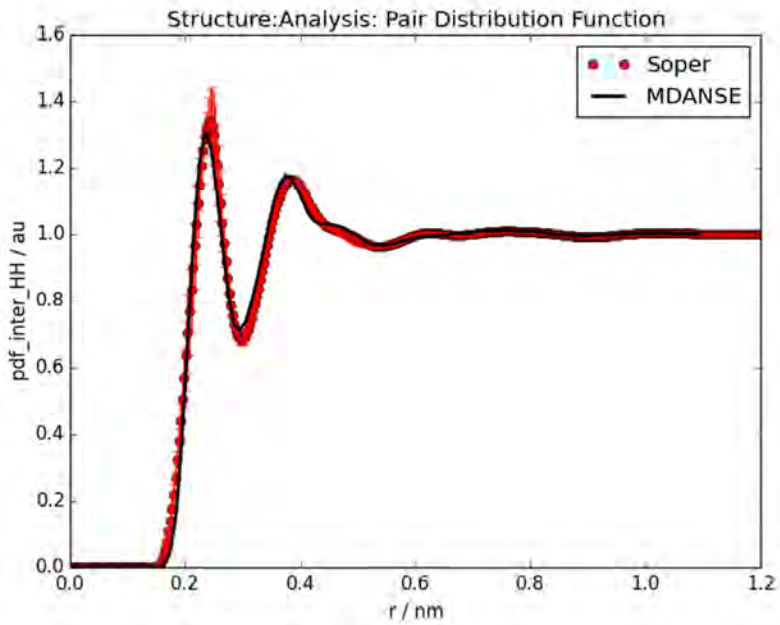


Fig. 9: Hydrogen - hydrogen pair distribution function: Simulation vs experiment.

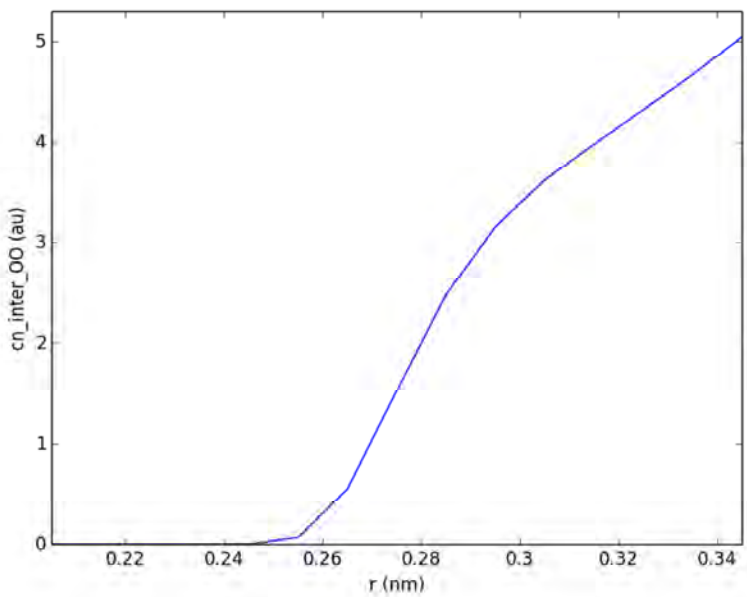


Fig. 10: Oxygen-oxygen coordination number. The value at the first minimum of the pdf (R_{min}≈0.33 nm) gives an indication of the number of molecules in the first coordination shell. In this case we have a coordination number slightly larger than 4, as expected for tetrahedrally coordinated liquid water.

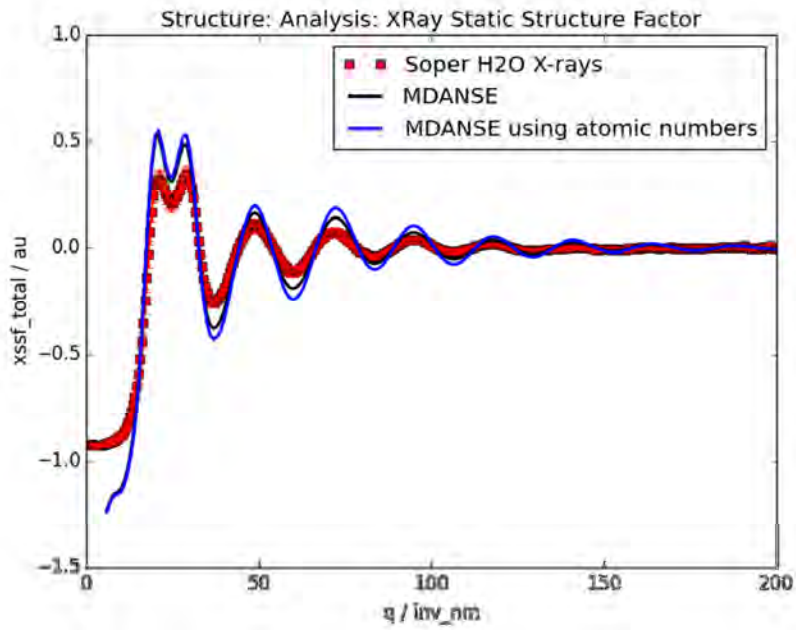


Fig. 11: Comparison of the experimental total static structure factor measured with x-rays and the simulated one with and without taking into account the atomic form factors.

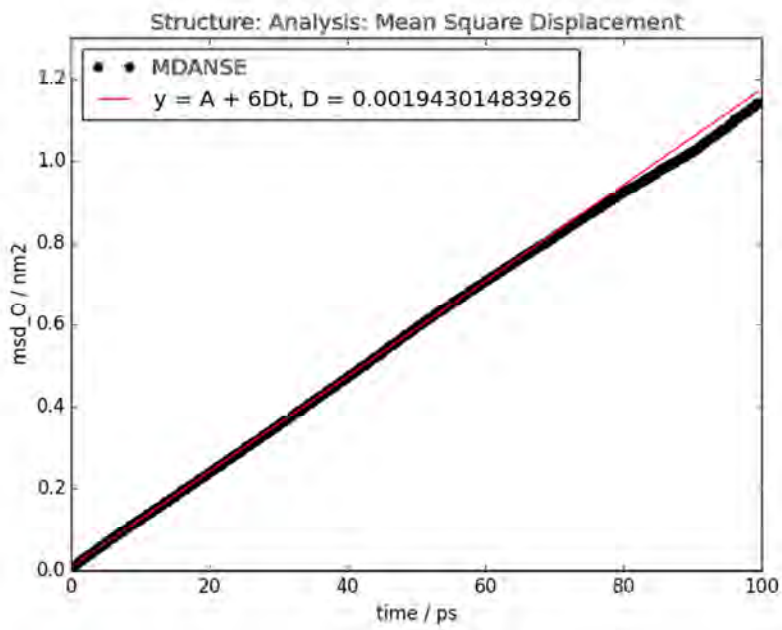


Fig. 12: Mean square displacement of the oxygen atoms and fit giving a self-diffusion coefficient D = 1.9×10^{-3} nm^2/ps = 1.9×10^{-9} m^2/s.

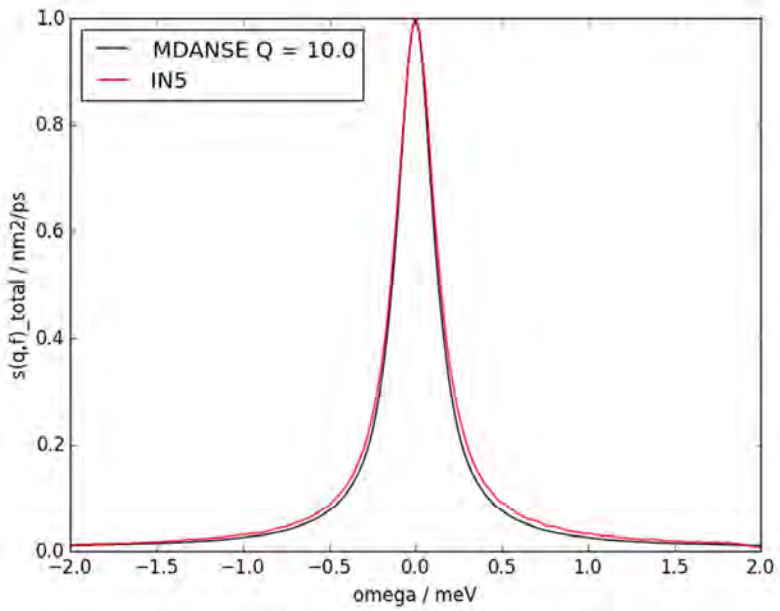


Fig. 13: Comparison of the computed dynamic incoherent structure factor for Q = 10 nm^{-1} with experimental data measured on IN5.

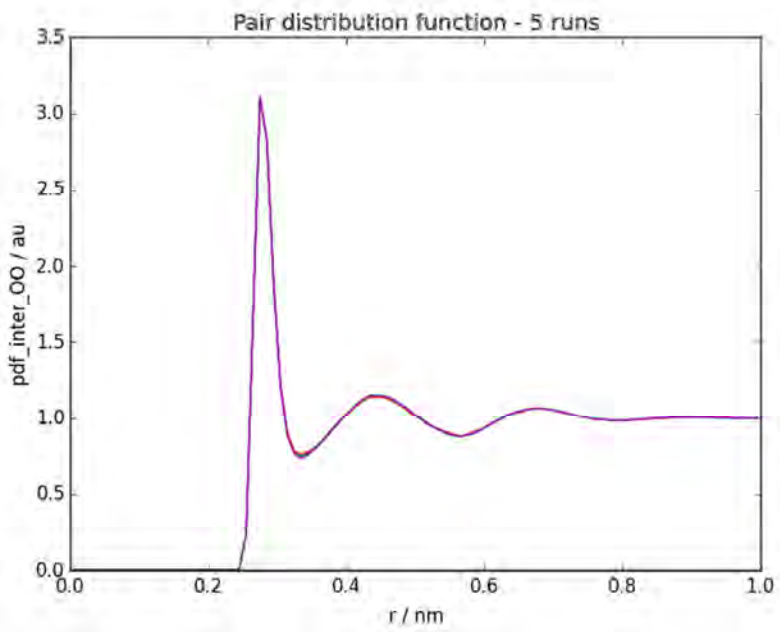


Fig. 14: Oxygen-oxygen p.d.f. computed for 5 independent trajectories. Each one is computed using 500 frames, so the errors are small enough to make the 5 curves almost indistinguishable.

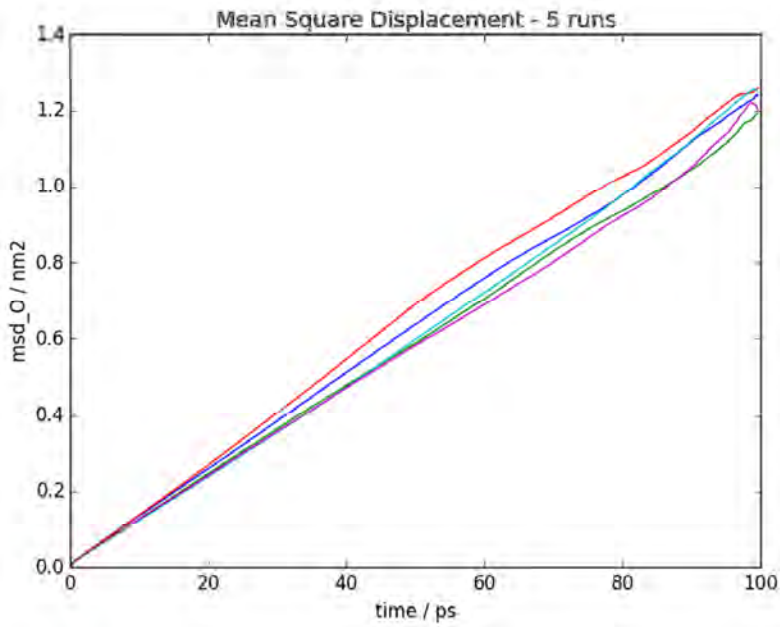


Fig. 15: Oxygen m.s.d. computed for 5 independent trajectories. In this case, the number of origins over which each point in time is computed decreases with time (from 500 for t=0 to only 1 for t = 99.8 ps), so the error bars will increase proportionally to $t^{½}$. The increase in the deviation between the 5 curves with increasing time shows clearly this effect.

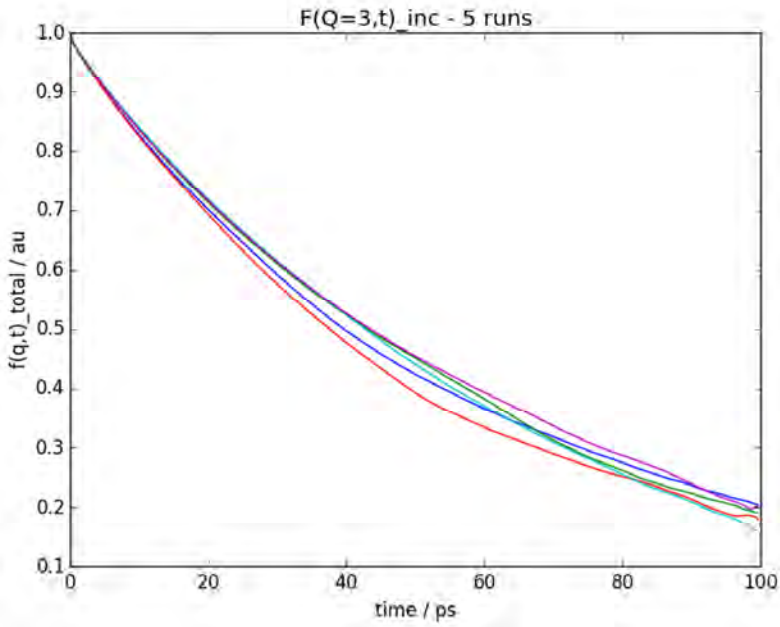


Fig. 16: Incoherent intermediate scattering function at Q = 3 nm^{-1} for 5 independent trajectories. Again the number of origins over which each point in time is computed decreases with time, making the 5 curves to deviate from each other more and more with time.

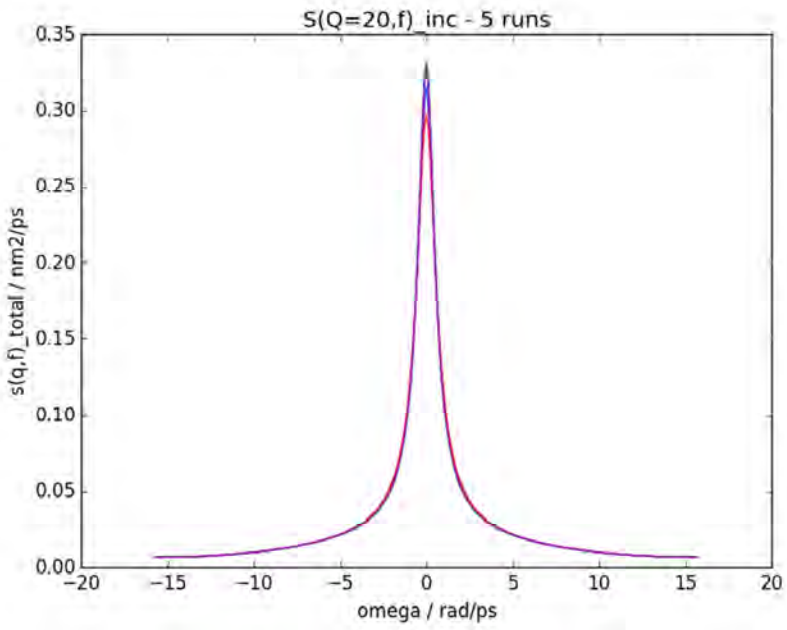


Fig. 17: Dynamic incoherent structure factor at Q = 20 nm^{-1} for 5 independent trajectories. Differences between the 5 runs are visible in the height of the elastic peak.

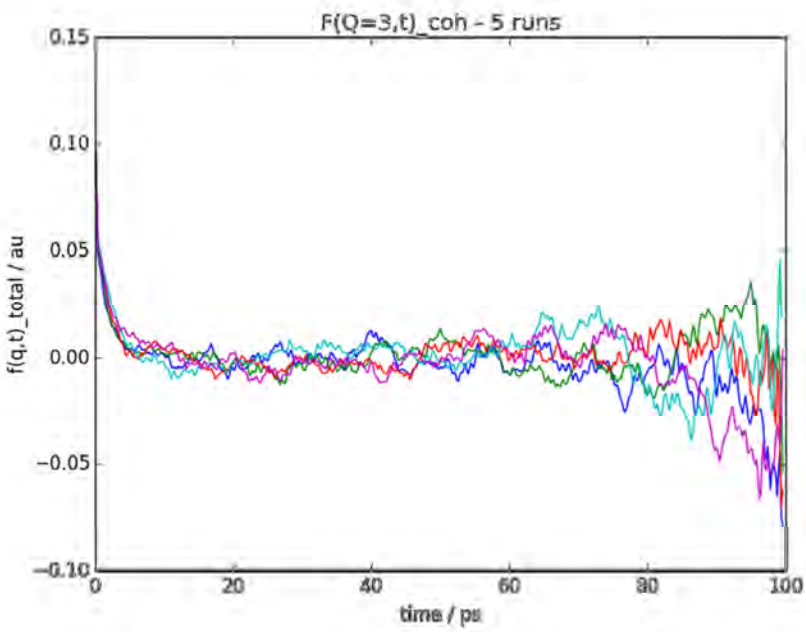


Fig. 18: Coherent intermediate scattering function at Q = 3 nm^{-1} for 5 independent trajectories.

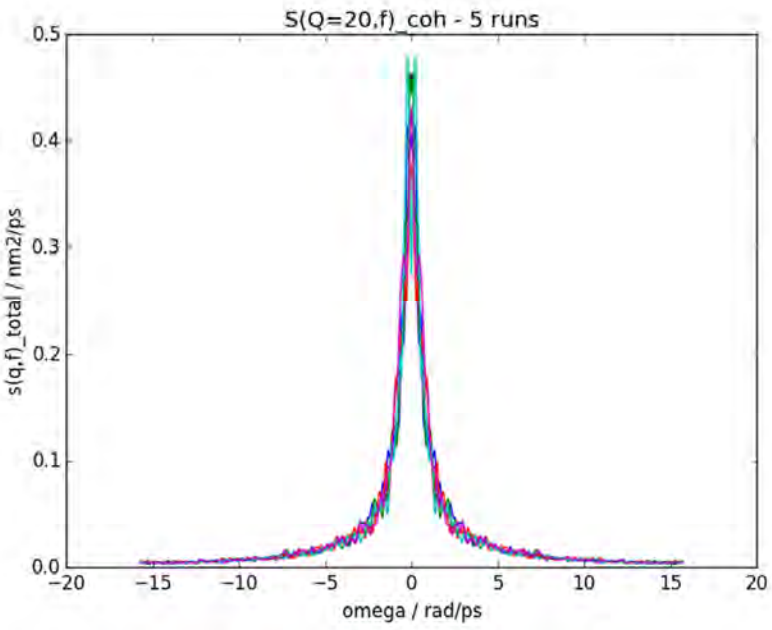


Fig. 19: Dynamic coherent structure factor at Q = 20 nm^{-1} for 5 independent trajectories.

Lattice Dynamics and INS Spectroscopy Tutorial

MDANSE Workshop 2016

Keith Refson and Sanghamitra Mukhopadhyay

Introduction

In this tutorial we are going to use *Materials Studio* program from Biovia to use CASTEP for lattice dynamics simulations to compare inelastic neutron scattering (INS) spectrum obtained from experiments.

We will use three molecular solids, NaH, benzene and toluene, for our calculations as they show distinctive INS features despite of their simple geometrical structures. Experimental INS spectrum are also provided to compare with the calculated one.

Vibrational frequencies of all these solids are calculated following the methodology normal mode analysis of lattice dynamics and within harmonic approximations.

The dynamic structure factors of INS spectroscopy are calculated from the normal mode frequencies and wave vectors following the equation:

$$S(Q,\omega) = \sum_n \sum_i^N \sigma_i \left\langle \frac{|Q \cdot U_{m,i}|^{2n}}{n!} e^{-2W_i(Q)} \delta(\hbar\omega - n\hbar\omega_k) \right\rangle_{\substack{\text{powder}\\ \text{average}}}$$

where **Q** is momentum transfer vector, ω_k is the ***k***-th lattice mode, ***U*** is the sum of all displacement vectors associated with lattice modes corresponding to ***i***-th atom having mass ***m*** and scattering cross section σ. The parameter ***n***=1 for transition (0 ⟶ 1), i.e., ground state to fundamentals, 2 for transition (0 ⟶2), i.e., ground state to second harmonic state and so on. The exponential term is the Debye Waller factor denoted by:

$$exp\{-2W_i(Q)\} = exp\Big\{ -|Q \cdot \sum_{m=1}^{3N-6} U'_{m,i}|^2 \Big\}$$

at low temperatures. At any temperature ***T*** , the full expression of the Debye Waller factor will be:

$$exp\{-2W_i(Q)\} = exp\Big\{ -\Big(Q \cdot \sum_{m=1}^{3N-6} \Big\{ u_i^\nu coth(\frac{\hbar\omega_\nu}{2k_B T}) \Big\}\Big)^2 \Big\}$$

where u_i^ν is the displacement vector associated with frequency *ν*.

Tutorial 1: CASTEP Phonon calculations for INS

The workshop PCs have been pre-configured with the *Materials Studio* program from Biovia which includes CASTEP and the graphical user interface to configure and drive the calculations.

Start Materials Studio

> You should see the main window, a set of toolbars, a “project explorer” panel on the left-hand side and depending on previous use) a prompt enquiring whether to open a new or existing project.

Create a new project. Unless you were presented with a prompt window for this you can do this by the menu "File"->"New Project"

> A pop-up window will ask you to choose a suitable name.

Choose any suitable name. This tutorial will assume "NaH-Phonon"

> You will see a "Project Explorer" on the left with the new project and a blank canvas on the right.

SETTING UP THE STRUCTURE of NaH

Now you must create a crystal structure model of NaH. There are various options for doing this

- Create a crystal structure from lattice parameters and co-ordinates. We will not do this here, but there is a Materials Studio tutorial to explain how to do this.
- Import an existing crystal structure

Menu "File->Import" will pop up a navigation menu. If you had a "cif" or other structure file you could import it from here. But Materials Studio also possesses a database of crystal structures, which can be found under "Structures".

NaH is not present in the database, but you will find NaCl. Locate and select NaCl.

> You will now see a “barebones” representation of cubic crystal structure of NaCl.

Bring up the "display style" menu using a right mouse click (or find it as one of the toolbar icons) and select Ball and Stick. Experiment with the mouse to rotate and scale the structure.

Now transform NaCl into NaH. First bring up the "Properties Explorer" using the menu "View"→"Explorers"→"Properties Explorer".

> You will see a long list of atomic attributes on the left below the "Project Explorer".

Finally you should change the name of the crystal from NaCl to NaH in the Project Explorer.

Click the model view to select one of the Cl atoms. You can now change this into H by double-clicking "Element Symbol" in the properties explorer.

> You will be presented with a periodic table to choose the new element. Choose H of course.

Make sure you have changed *all* of the Cl into H. Hint - use the right mouse menu to find all of the symmetry images.

> You ought now to have a NaH structure but with the lattice parameter of NaCl.

For calculation purposes it is best to work with the primitive cell. You can change to this using the "Build"->"Symmetry" menu.

The cubic conventional cell view should be replaced by a primitive cell one.

We could simply edit the lattice parameters, but instead let CASTEP do the work.

Setting up a CASTEP Geometry Optimisation

Start the CASTEP setup driver - the "wavy lines" icon on the tool bar.

You will see a window appear labelled "CASTEP Calculation"

There are several choices and options to set. Under the "Setup" tab choose:

Task: Geometry Optimization
Click "More" and choose "**Optimize cell**"
Exchange-Correlation Functional "PPE-GGA"
Quality: Fine - this chooses the "quality" and size of the basis set.
Metal: You should **uncheck** this box (otherwise LO/TO splitting will not be calculated in the subsequent phonon calculation).

Under "Electronic"

Pseudopotentials: "Norm Conserving"
k-point set: "Medium" (For speed)

Job Control
Choose the maximum number of CPU cores allowed.

Now you are ready to run the calculation. Click "Run"
If you did not already choose it, you will be asked to choose whether to convert to use a primitive cell for this calculation. This is quicker so check "Yes"

The job will be submitted to SCARF but should run and complete quickly - around 30s.

You will be presented with a window containing the CASTEP output file. This contains several geometry optimization steps (BFGS iterations) for a constant-pressure optimisation.

Near the end of the file you will find the "BFGS: Final Configuration" and a description of the final lattice parameters.

What is the cubic lattice parameter of NaH? Hint - you will need some arithmetic to work this out. How does it compare with the experimental value of 4.89

Analysis of Geometry Optimization

To get the optimised structure back into Materials Studio you must start the "CASTEP Analysis" tool in Materials Studio. You will find this under the "Wavy lines" tool icon.

The analysis window will pop up giving a list of things to analyse.

From the list choose "Structure". To allow the import it is essential that you first click the crystal structure display window. When you have done this, the “Update button will be active and you can click it.

At this point you can also import the electron density if you like.
You should see an electron density isosurface displayed.

Setting up the phonon calculation.

In the project explorer select the NaH crystal structure which should contain the updated structure. As before, start the CASTEP setup tool.

Choose the settings The choices you made for the geometry should be retained
Under "Setup"

Task: Energy (This will enable phonons and much more)
Quality: Fine - this chooses the "quality" and size of the basis set.
Metal: You should uncheck this box otherwise LO/TO splitting will not be calculated

Under "Electronic"

Pseudopotentials: "Norm Conserving"
k-point set: "Coarse" or "Medium" (For speed)

Under "Properties" Check the "Phonons" box..

You should be presented with a choice of DOS, dispersion or both. Both is good!

Click the "More" button. Then click the dispersion "Path" button to define the Q-space path through the Brillouin Zone.

A labelled Brillouin Zone should appear on top of your structure. Materials Studio sets up a good default path and there is no need to choose it.

*Theory note: The calculation does not compute the frequencies at each wavevector independently. Instead it computes the full force constant matrix of Lattice Dynamics theory using a regular grid of q-points, and uses Fourier methods to compute the DOS and dispersion. For this first run you should set the ""q-vector grid spacing for interpolation" to around 0.08. ** Warning - you can make the calculation arbitrarily expensive using this control!*

Now you are finally in a position to run the calculation. Click the "Run" button and wait….

Analysing the phonon runs

Under the "Wavy lines" toolbar item icon select " CASTEP Analysis".
You will see a list of possible analyses. Select "Phonon density of states"

To enable the reading in of the data you must first select the CASTEP output "NaH_PhonDOS" in the *Project Explorer* in the left-hand panel.
When you have done the correct selections, the "Results File" box in the Analysis window

should read "NaH_PhonDOS.castep"

You will then be able to click the "View" button
You should see a plot of the DFT calculated phonon density of states of sodium hydride.

Repeat the procedure but for phonon dispersion, being sure to select the correct output file.
You should see a plot of the phonon dispersion of sodium hydride. Congratulations!

You can also calculate some thermochemical properties by performing the "Thermodynamic Properties" analysis of the DOS results.
The result will be plots of energy, heat capacity entropy and Debye temperature as a function of T.
N.B. You have been provided with the files from the NaH phonon calculation ready to analyse, in case of computer troubles.

Tutorial 2: Molecular and Crystalline Benzene.

This tutorial compares calculations for a molecule of Benzene with the crystalline form, to illustrate how important crystal packing effects are to the vibrational spectra.

CASTEP always uses periodic crystal boundary conditions, so we will use a trick to simulate an isolated molecule by generating a crystal with a large lattice spacing which introduces a substantial vacuum space between molecules. (The supercell should not be too large though, as in CASTEP's plane-wave basis set, vacuum costs CPU cycles as well).

Geometry optimisation of the Benzene molecule

You have been provided with an input file of a benzene molecule in a 8Åx8Åx4Å hexagonal cell, named "benzene-mol-hex-c.pdb". To load this into Materials Studio the steps are:

1. Create a new project named, e.g. "molecular-benzene" from the File menu
2. Again from the "File" menu choose "import→structure", find and click on "enzene-mol-hex-c.pdb" on the memory stick.

CASTEP makes use of crystal symmetry to speed the calculation so we will need to add this using the "Build" menu.

1. Change the display type to ball-and-stick as you did for NaH.
2. Add a representation of the bonds using "Build->bonds"
3. Select "Build→Crystal" from the menu to add the unit cell.
4. Select "Build→Symmetry-Find Symmetry" to analyse the space group symmetry of the structure
5. Click the "Impose Symmetry" button"

You should now have a benzene molecule ready to set up a CASTEP calculation. Start the CASTEP run tool and set up the geometry optimization

Task: Geometry Optimization
This time DO NOT select "**Optimize cell**"

Exchange-Correlation Functional “PPE-GGA”
DFT-D correction: Select the “TS” dispersion correction
Quality: Fine - this chooses the "quality" and size of the basis set.
Metal: You should **uncheck** this box (otherwise LO/TO splitting will not be calculated in the subsequent phonon calculation).

Under "Electronic"

Pseudopotentials: "Norm Conserving"
k-point set: "Gamma" IMPORTANT: This is different from NaH.

Job Control
Choose the maximum number of CPU cores allowed.

From here the calculation proceeds as for NaH,

Phonon calculation of molecular benzene

Setting up and running the phonon calculation is very similar to NaH *except*

You need to find the “More” button for phonon properties calculation and

1. uncheck the box marked “use interpolation”.
2. set the Phonon DOS quality to “Gamma”

This instructs CASTEP to perform a single phonon calculation for q=0 in contrast to the Brillouin-Zone sampling needed for a crystal.

Phonon analysis of molecular benzene

This proceeds in a similar fashion to NaH.

I suggest you carefully read the main text output which pops up in a window, and find the block containing the output frequencies and group-theory symmetry analysis, plus any ir and Raman intensities.

Optional – IR and Raman spectra

Try repeating the phonon calculation but additionally check the box “Polarisability, IR and Raman” in the *properties* selection. The run will take several times longer if you check the tick box for Raman intensities too, as this is currently an expensive calculation.

To display the IR spectrum you will need to

1. double click to select the molecular structure display window
2. in the analysis tool, choose the property “IR spectrum” and click “Import the Hessian”

Analyse crystalline benzene

A phonon calculation of crystalline benzene is supplied on your USB stick. You should be able to navigate within Materials Studio to “Open Project”.

As with the molecular calculation, examine the text output to find the frequencies, group-theory analysis etc.

Compare the DOS with the molecular case.

Toluene

Similar to benzene a set of calculations on Toluene has been provided on the USB stick.

Using a-climax to compute INS spectrum from CASTEP calculations

aClimax is the software for calculations of INS (inelastic neutron scattering) spectra from the output of phonon calculations using CASTEP, GAUSSIAN, CRYSTAL16. The software is designed to produce INS spectra at low temperature (T~ 10K).

Start a-Climax by double clicking a-Climax application as provided.

On the aClimax window click on Files -> LoadData

A window will appear to load phonon data. Select NaH_PhonDOS to Open. This file is in :

C:\MDANSE 2016\LatticeDynamicsINS\SimulationsData\NaH\NaH_Files\Documents\NaH CASTEP GeomOpt\NaH CASTEP Energy (2)\ directory.

A window will appear asking whether to load a new file. Keep the default option “Identify D automatically” checked on. Click on **OK**.

Another window will appear giving “Wing Status”. Move the window box to the right side of the application box without changing anything.

In the leftmost window box named as “Main Interface”, unchecked box “Include Wing Calculations” and also “Perform Smoothing of Wings”. The concept of ‘wings’ appear when one wants to approximate solid state INS spectra from molecular level simulations. When we do simulations on periodic solid then no wing has to be considered.

Click on the button called “**INS Calculations**” on “Main Interface” window.

When INS calculation will finish a dialogue box informing “Done” will appear. Click on “OK”.

On the main aClimax application top bar, click on Windows -> Spectrum

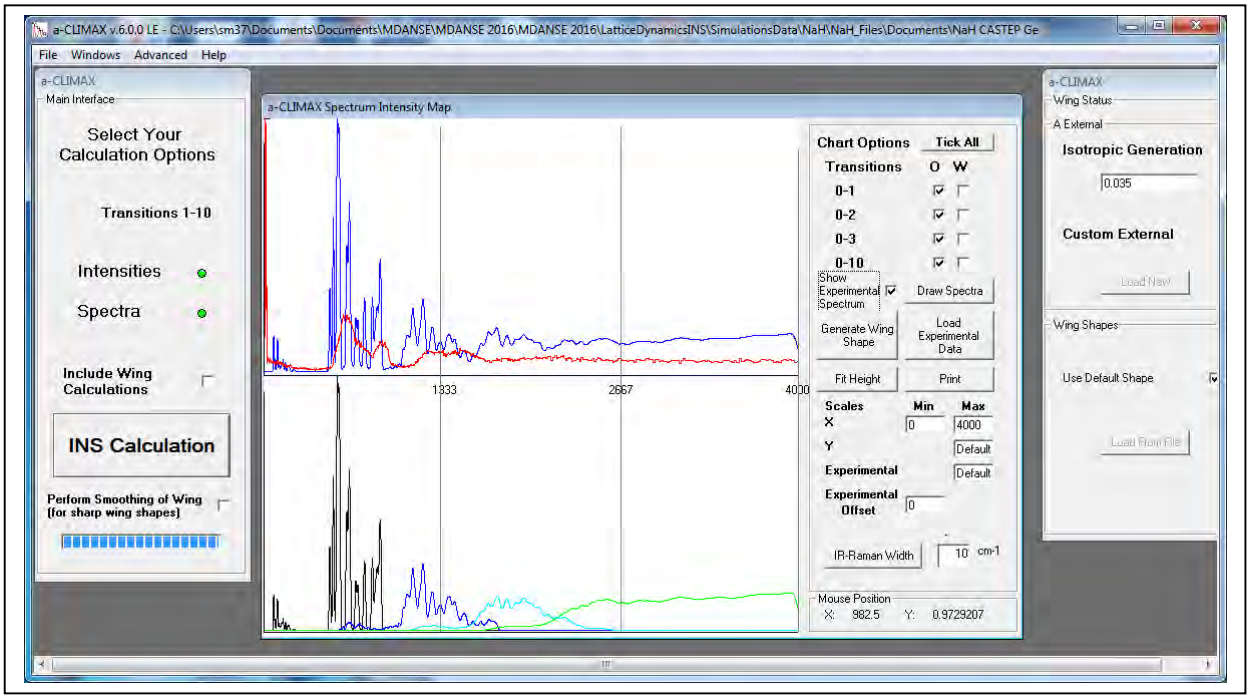
Another window named “Spectrum Intensity Map” will appear. On the left hand side of that window several buttons and check box options are available. Study those options carefully. The 0-1, 0-2, …0-10 indicate fundamentals, first overtone, second overtones,..etc. upto 10th order overtones. First click on the check box below “O” for 0-1 transition. The “O” stands for overtone. Click on the button of “Draw Spectra”. The INS spectra for fundamentals (0-1 transition) will appear on the display region of that window.

Click on for other transitions, such as 0-2, 0-3 etc. and click on “draw spectra” for the updated spectra to visualize.

The top panel of the visualization window shows the total INS spectra where the bottom panel will show the contributions from individual transitions in different colours.

To load experimental spectra click on the “Load Experimental Spectra” button. The input dialogue box will open. Select any of the NaH experimental file has been provided within \LatticeDynamics\ExperimentalData\NaH\ directory. Check on “Show Experimental spectra” check box. Click on “Draw Spectra”. The experimental spectra will appear along with the simulated INS spectra.

To save this spectra as Ascii file, go to the top option panel of aClimax main window. Go to File -> Export Spectrum. In the dialogue box, select CSV file to save the plotted data.



Use the same procedure to calculate INS for benzene and toluene and compare with experimental data.

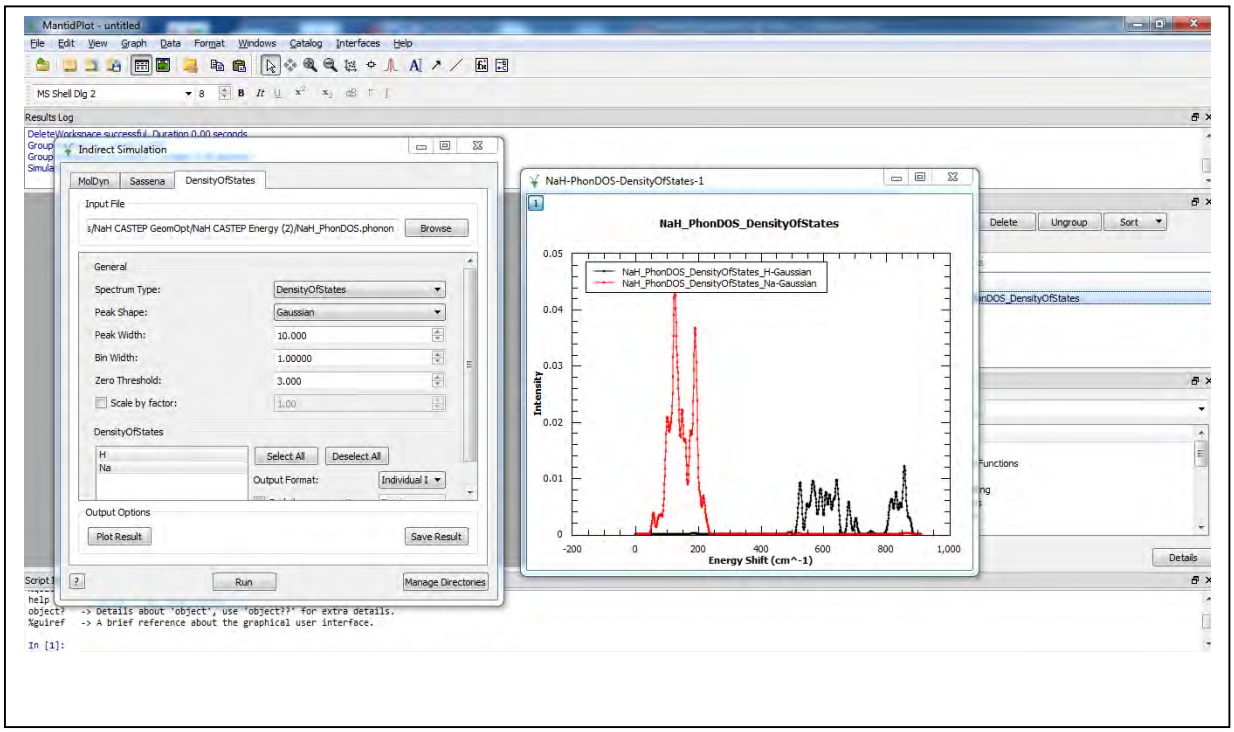
What is the difference between molecular level and solid state level calculated INS spectra ?

Using Mantid to analyse phonon VDOS

Mantid is the data analysis tool for neutron spectroscopy.

Start Mantid. Set up directory to save data by browsing to the preferred directory through File -> Manage Users Save Directory -> Default Save Directory.

Load interface by clicking Interfaces -> Indirect -> Simulations and select the DensityOfStates option.



Load NaH_PhonDOS.phonon as input file. Keep all options as default and click on **Run**. The calculated file will appear on the workspace. Right click on the file. Select Plot spectrum option. In the appeared dialogue select spectrum number 0 to plot the VDOS. The VDOS of both Na and H will be plotted.

To get the position of the vibrational density as stick plot select spectrum 1.

The peak width can be changed to adjust the resolution of the instrument.

To weight the individual atomic peak with their neutron cross section click on scale by cross section on the original interface dialogue.

Results can be saved in Ascii format by clicking Save -> Ascii at the top bar menu at the top of the workspace window.

What is the difference in VDOS when you select NaH_PhonDisp ? Why?

References

1. R. M. Martin, *Electronic Structure, Basic Theory and Practical Methods,* Cambridge University Press, (2004)

2. M. T. Dove, *Introduction to Lattice Dynamics*, Cambridge University Press (2005).

3. S. W. Lovesey, *Theory of Neutron Scattering from Condensed* Matter, Oxford University Press (1984).

4. P. C. H. Mitchell et al, *Vibrational Spectroscopy with Neutrons*, World Scientific (2005).

5. G. J. Kearley, Nuclear. Inst. Method. in Phys. Res. A, **354**, 53 (1995).

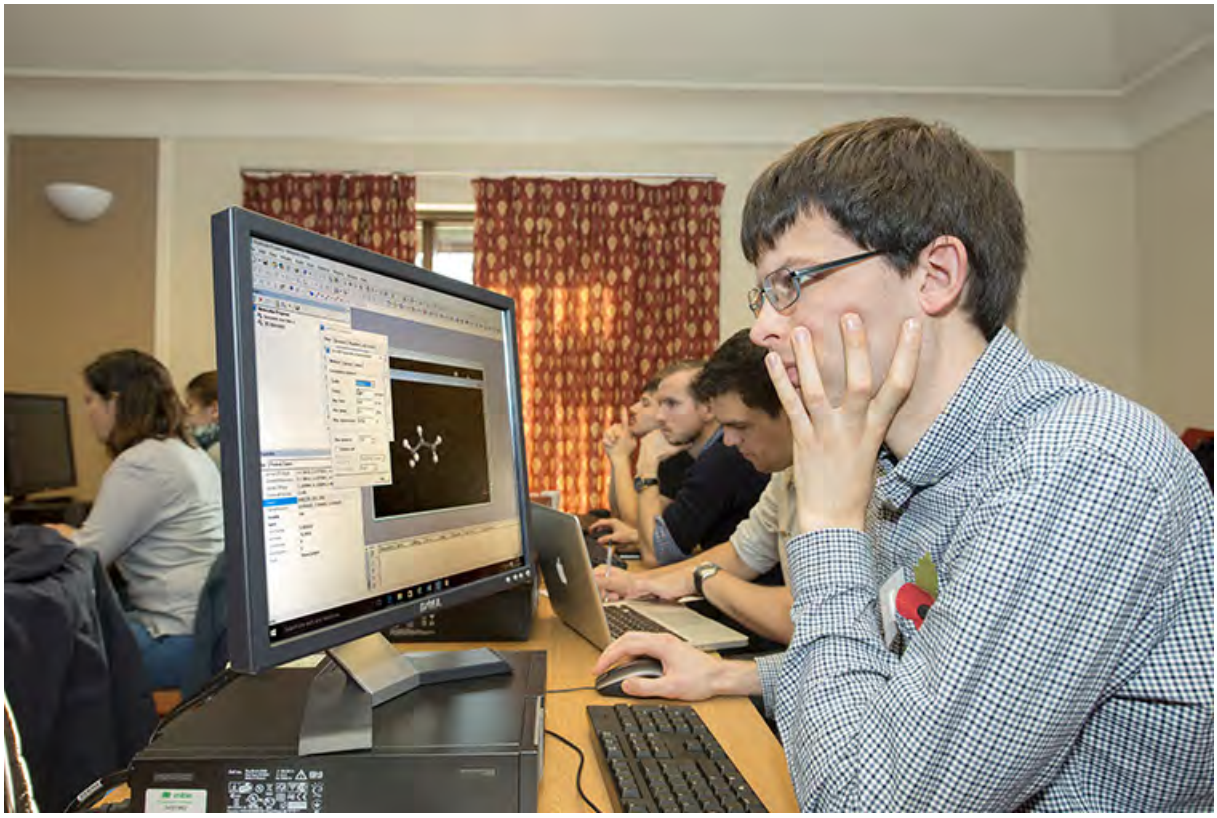
6. D. Champion, J. Tomkinson, G. Kearley, Appl. Phys. A, **74**[Suppl.] S1302 (2002).

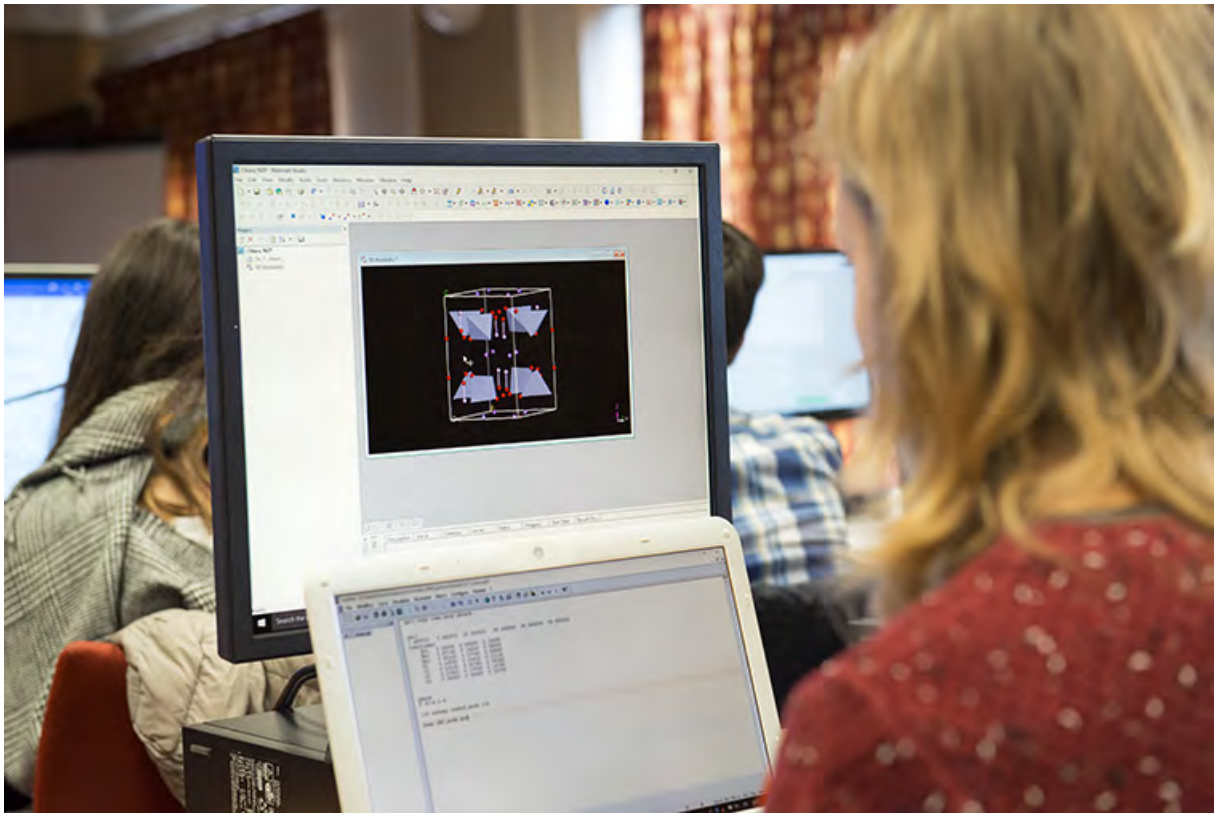
7. G. L. Squires, *Introduction to the Theory of Thermal Neutron Scattering*, Dover Publications (1996).

8. A. J. Ranirez-Cuesta, Comp. Phys. Commun., **157**. 226 (2004).

9. http://www.castep.org/

10. http://www.mantidproject.org/Main_Page





DELL