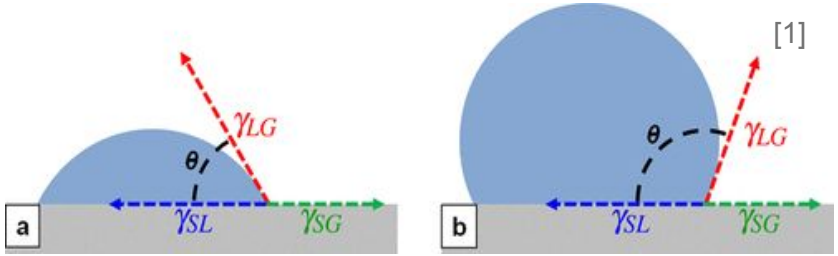
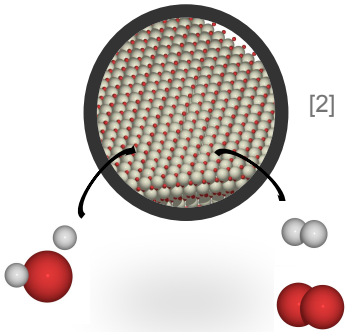
LAB: Predicting vibrational frequencies from the external chemical environment: OH at hydrated metal-oxide surfaces

Jolla Kullgren
Department of Chemistry - Ångström

Background



[1]

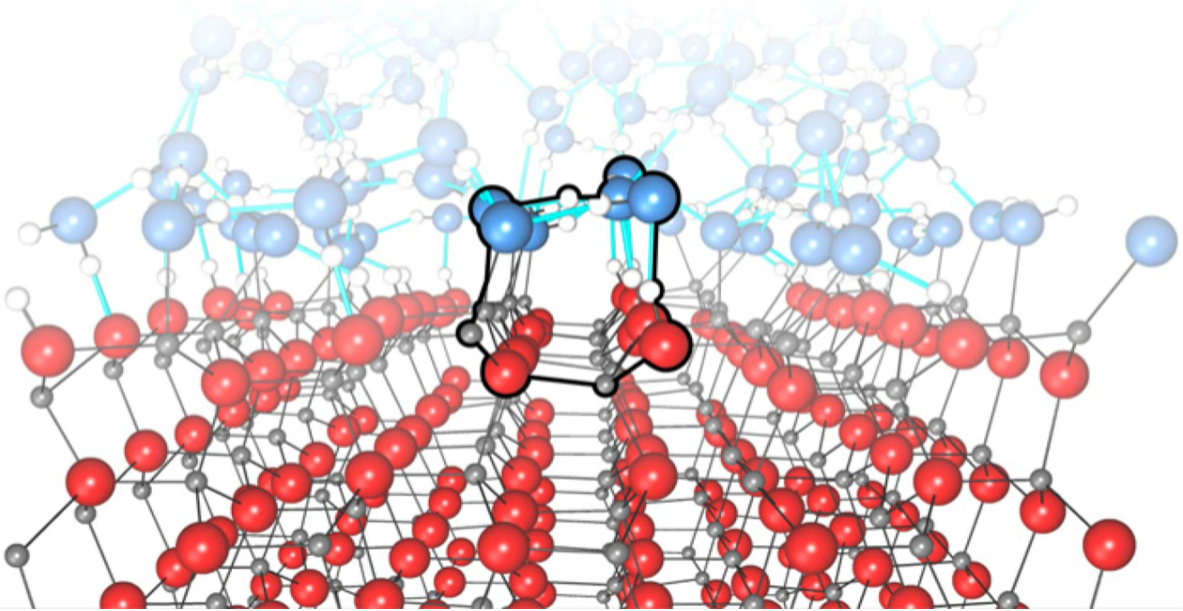


[2]

- The surface water affect:
- Hydrophobic/hydrophilic properties
- Regulate adsorption of biological molecules
- Surface-driven reactions

[1] Carchini et al. ACS Appl. Mater. Interfaces, vol. 8, no. 1, pp. 152–160, 2016.
[2] Rodriguez Chem. Soc. Rev., vol. 46, no. 7, pp. 1824–1841, 2017.

The surface/water interface



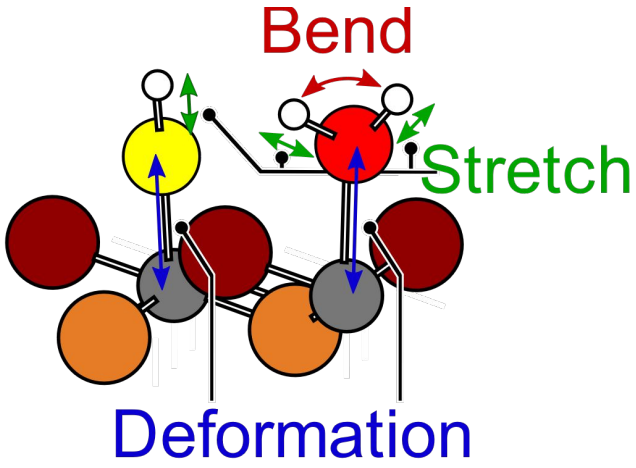
Water molecules close to the metal oxide surface often engage in structural motifs quite different from that of bulk water. These structures are very difficult to directly observe with experimental techniques (only in the case of low water coverage, low temperatures and high vacuum conditions).

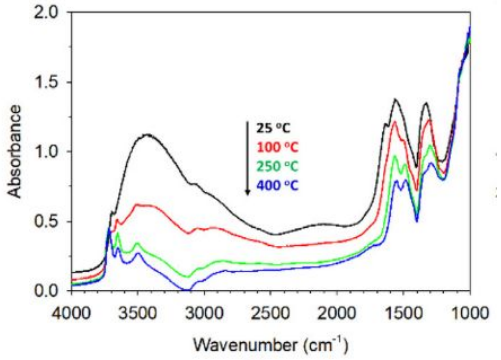
UPPSALA UNIVERSITET

Figure: Raymand D. (2010). Uppsala University, Uppsala.

The surface/water interface

Signals from vibrational spectroscopy could be used to infer the structure, at least partially, provided we have a model/approach to make the assignment. Or reversely, ***what information from the external chemical environment surrounding the oscillating molecule is required to accurately predict it's frequency?***





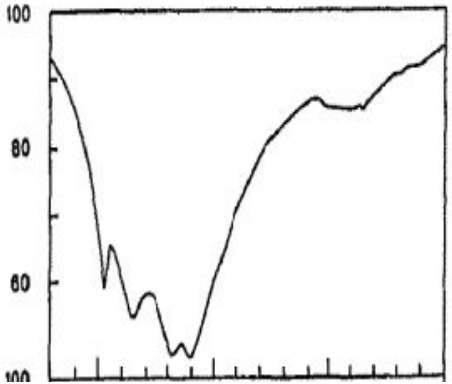
Luo et al. Surf. Sci., vol. 691, pp. 0039–6028, Jan. 2020.

UPPSALA UNIVERSITET

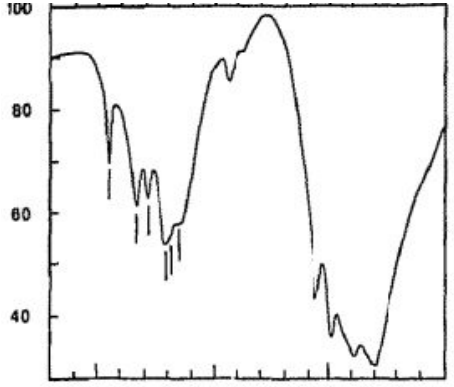
Partitioning the world

The question only makes sense if we can decouple the vibration of the molecule from it's surroundings. Experimentally, we can make use of isotope isolation (here: replacing hydrogen with deuterium).

100 % H_2O



5 % H_2O 95 % D_2O



91,

).

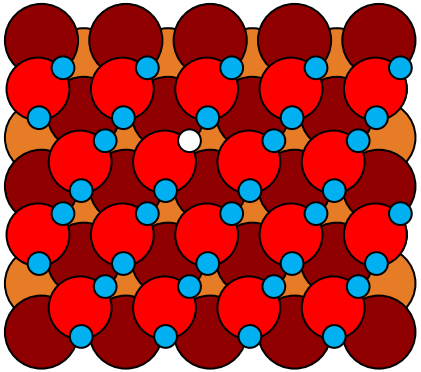
UPPSALA UNIVERSITET

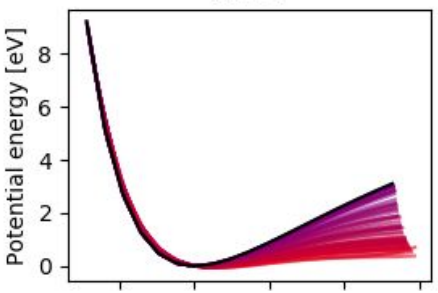
[1] Berglund, Lindgren, Tegenfeldt, *J. Mol. Struct.*, vol. 43, no. 2, pp. 169–177, 1978.

Partitioning the world

Question only makes sense if we can decouple the vibration of the molecule from it's surroundings. Computationally, we can do something similar. We consider the potential energy surface related to the stretching of a single O-H bond in our system.

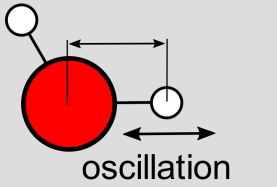
Red= Oxygen, White=Hydrogen, Blue=Deuterium, Brown="The surface"

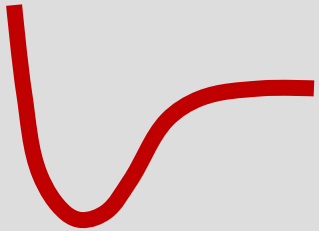




Discrete variable representation (DVR)

Generate PES
OH stretch





Construct Hamiltonian using basis functions

$$H = T^T \cdot T^{fbr} \cdot T + V^{PES}$$

$$T^{fbr} = \frac{1}{2\mu}\begin{pmatrix} t_1^{fbr} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & t_n^{fbr} \end{pmatrix}$$

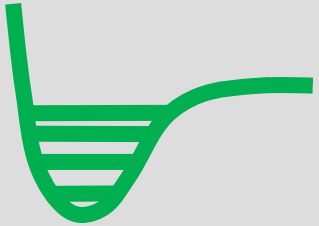
$$t_n^{fbr} = \left(\frac{n\pi}{x_{max} - x_{min}}\right)^2$$

$$T = \sqrt{\frac{2}{N+1}}\begin{pmatrix} t_{1,1} & \cdots & t_{1,j} \\ \vdots & \ddots & \vdots \\ t_{i,1} & \cdots & t_{i,j} \end{pmatrix}$$

$$t_{i,j} = \sin\left(\frac{(i+1)(j+1)\pi}{N+1}\right)$$

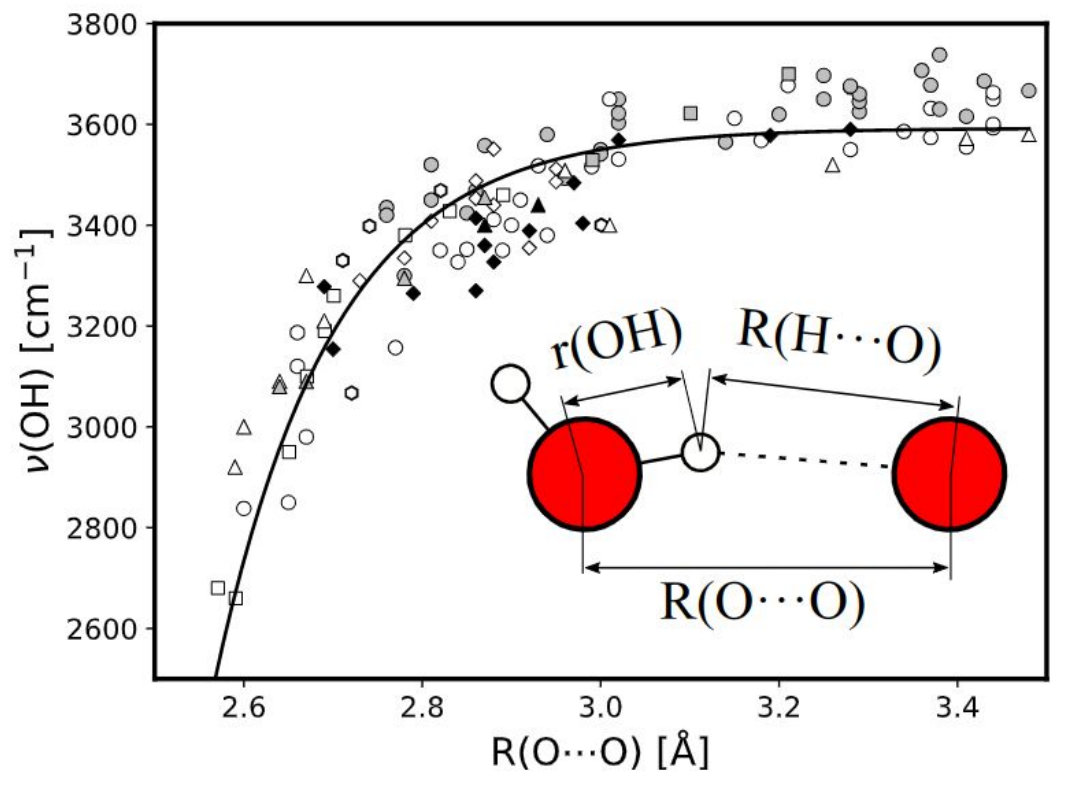
Diagonalize for eigenvalues and eigenvectors

Frequencies from ΔE



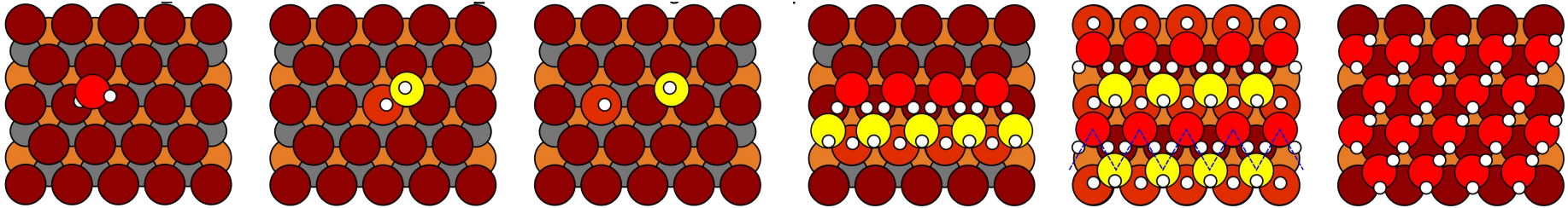
[1] Light et al J. Chem. Phys., vol. 82, no. 3, pp. 1400–1409, 1985.
[2] Bacic. Annu. Rev. Phys. Chem., vol. 40, no. 1, pp. 469–498, 1989.

Experimental bulk correlations



Adopted from: E. Libowitzky, Monatshefte für Chemie, 130, 1047 (1999)

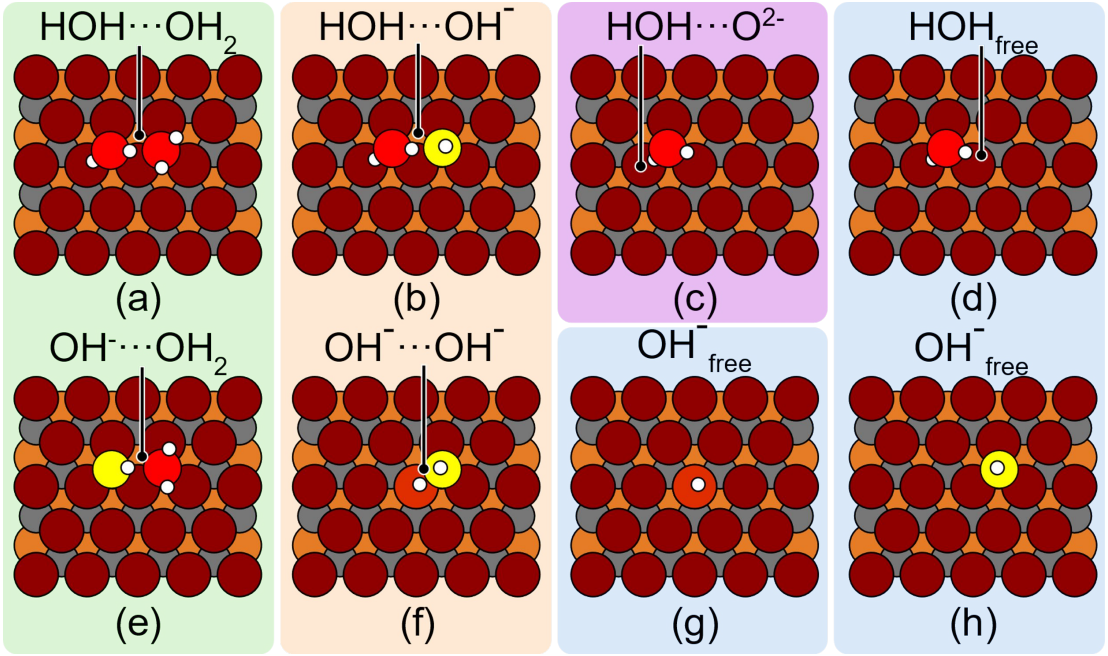
Data-set: Systems



Surface	Water coverages	Total no. of structurally unique water molecules (intact or dissociated)	No. of unique OH groups investigated		
			O_wH	O_sH	OH_f
CeO_2(111)	1-,2-mers, 1ML(many) 1.5ML	65	68	29	29
MgO(001)	1-,2-,3-mers, 1ML,1.25ML	12	16	4	4
CaO(001)	1-,2-,3-mers, 1ML,1.25ML	34	41	13	13

UPPSALA UNIVERSITET

Data-set: Motifs

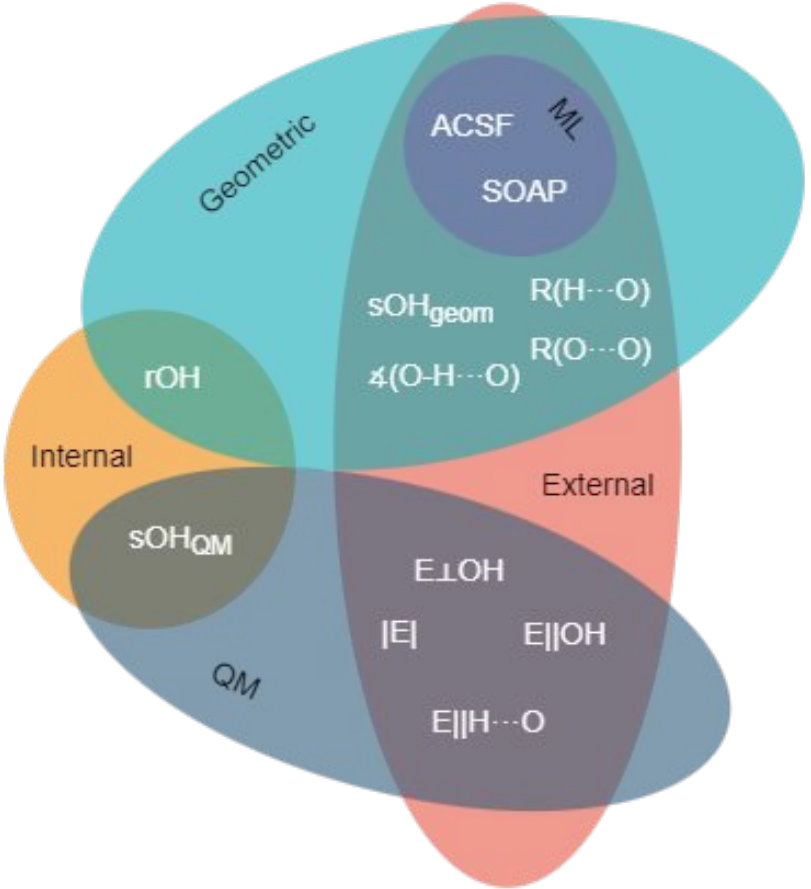


Acceptor	Frequency range [cm^{-1}]
free	3600 - 3450
H_2O	3500 - 3300
OH^-	3500 - 2350
O^{2-}	3300 - 2500

The range is largely dependent on the acceptor.

Data-set: Descriptors

- The data-set contains a large number of descriptors
- The external geometrical descriptors are the typical descriptors used in the experimental literature.
- External QM descriptors comprise of the electric field calculated with our DFT-approach.
- sOH refer to the so-called bond-order.



The assignment

The context is “Predicting vibrational frequencies from the external chemical environment”. That is given a number descriptors, what can we infer about the frequency. Or in reverse, what information regarding the structure is hidden in the frequency. You can use the proposed work-flow below, or explore the data more freely. However, you should demonstrate the use of the techniques we have covered in the lecture/workshops.

Proposed work-flow:

- Load the csv-files with with frequencies and descriptors into a pandas data-frame.
  - NOTE: There is a reference bulk data-set for comparison alongside the surface data described above.
- Perform a few ANOVA tests for all/selected descriptors in the surface and bulk data-set to get a feel for what correlation there are among the various descriptors.
- Perform PCA or LDA analysis on the data. E.g. to compare if bulk and surface data can be discriminated.
- Perform a number regressions and try to find a model with a reduced number of independent variables that can make robust prediction of frequencies.
- Make a correlation plot similar to that of Libowitzky (slide 8), but with surface and bulk data highlighted. Draw a fitted model curve from e.g. i kernel ridge regression on top of the data.
- Summarise your findings into a reasonably comprehensive jupyter notebook. Write down some reflections and interesting findings from your “research”.