ΦSX

Quantum Mechanics of the Chemical Bond

Peter E. Blöchl

Caution! This is an unfinished draft version
Errors are unavoidable!

Institute of Theoretical Physics; Clausthal University of Technology;
D-38678 Clausthal Zellerfeld; Germany;
`http://www.pt.tu-clausthal.de/atp/`

don't panic!

© Peter Blöchl, 2000-February 22, 2021
Source: `http://www2.pt.tu-clausthal.de/atp/phisx.html`

Permission to make digital or hard copies of this work or portions thereof for personal or classroom use is granted provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation. To copy otherwise requires prior specific permission by the author.

[1]

---

[1] To the title page: What is the meaning of ΦSX? Firstly, it sounds like "Physics". Secondly, the symbols stand for the three main pillars of theoretical physics: "X" is the symbol for the coordinate of a particle and represents Classical Physics. "Φ" is the symbol for the wave function and represents Quantum Physics "S" is the symbol for the entropy and represents Statistical Physics.

Foreword and outlook

This book shall provide some chemical insight for physicist and material scientists. This book is devoted to simple concepts that only require back-on-the-envelope calculations. The emphasis is on qualitative considerations instead of accurate predictions. The book attempts to provide a structure into the thinking about materials properties and reactive processes.

The for this course was born out the the experience with students venturing into ab-initio simulations. What I found necessary is first a good education in basic theoretical physics and secondly an intuitive feel for materials. These book addresses the second aspect.

There have been many attempts to condense quantum mechanics into simple concepts that allow one to understand materials without difficult mathematics and computer simulations. The most important contribution is probably due to Linus Pauling's book "The Nature of the Chemical Bond", which was written already in 1939. His concepts are still large part of a lecture on structural chemistry. Later contributions are due to Walter Harrison, O.K. Andersen and Roald Hoffman and many more.

Simple qualitative arguments are important for a computer scientist to ask the right questions, and to analyze the results in a way that practitioners can understand. They are also useful to cross-check the results of computer simulations. Calculations are never perfect. Problems, even ones as trivial as setting up the incorrect starting structure, tend to plague the practitioner. Having an idea of what is likely to happen allows one to detect and correct problems.

The experimentalist in the lab has very similar problems. He has to interpret his results, understand side effects, and detect if he has made an important observation, that brings him the Nobel price, or simply a bad experiment.

The second goal of this book is to promote the understanding among experimentalist and theoreticians on the one hand and physicists, chemists, biochemists and materials scientists on the other.

Books[1]

Contents

1 Brief review of quantum mechanics 11
- 1.1 Classical mechanics . . . 11
- 1.2 Quantum Mechanics . . . 11
  - 1.2.1 Hamilton operator and observables . . . 11
  - 1.2.2 Schrödinger equation . . . 12
  - 1.2.3 Expectation values . . . 12
  - 1.2.4 Bracket notation . . . 13
- 1.3 Many non-interacting electrons and Pauli principle . . . 14

2 The covalent bond 15
- 2.1 Linear combination of atomic orbitals (LCAO) . . . 15
- 2.2 From non-orthonormal to orthonormal basis functions . . . 16
- 2.3 The two-center bond . . . 18
- 2.4 Bonds and occupations . . . 22
- 2.5 Three-center bond . . . 23
- 2.6 Chains and rings . . . 25
  - 2.6.1 Chains . . . 25
  - 2.6.2 Rings . . . 29

3 The use of symmetry 33
- 3.1 Introduction . . . 33
- 3.2 Symmetry and quantum mechanics . . . 34
- 3.3 Symmetry eigenstates of the hydrogen molecule . . . 36
- 3.4 Symmetry eigenstates of a main group dimer . . . 37
  - 3.4.1 Approximate diagonalization . . . 40
  - 3.4.2 σ bonds and antibonds . . . 40
  - 3.4.3 π bonds and antibonds . . . 41
- 3.5 Parameterized Hamiltonian of the main group dimer . . . 43
  - 3.5.1 Notion of σ, π, δ-states . . . 43
  - 3.5.2 Slater-Koster tables . . . 43
  - 3.5.3 Hopping matrix elements of Harrison . . . 43
  - 3.5.4 Parameterized Hamiltonian of the main group dimer . . . 44
  - 3.5.5 Diagonalize the sub-blocks . . . 46
- 3.6 Stability of main group dimers . . . 47
- 3.7 Bond types . . . 48
- 3.8 Worked example: orbitals of the ethene . . . 48

3.9 Worked example: orbitals of the methane . . . 52
3.10 Worked example: orbitals of the allyl ion . . . 54
3.11 Worked example: orbitals of closo-$B_6H_6^{2-}$ . . . 57
3.12 Exercise: sketch orbitals . . . 60

4 Hybrid orbitals 61
4.1 Spherical and cubic harmonics . . . 61
4.2 Rotation of p-orbitals . . . 63
4.3 Superposition of s- and p-orbitals: Hybrid orbitals . . . 64
4.4 Hybrid orbitals and structure . . . 66
4.4.1 Hydrides of main group atoms . . . 69
4.5 Coordination of transition metal compounds . . . 72
4.5.1 Square-planar coordination . . . 72
4.5.2 Tetrahedral coordination . . . 72
4.5.3 Octahedral coordination . . . 74
4.5.4 s-p-d hybrids . . . 74

5 Frontier orbitals 75
5.1 Determine strong bonds . . . 76
5.1.1 Nitronium ion . . . 76
5.1.2 Naphthalene . . . 78
5.1.3 Nitridation of Naphtalene . . . 79

6 From bonds to bands 81
6.1 The Jellium model . . . 81
6.2 Density of States . . . 82
6.2.1 Motivation . . . 82
6.2.2 Density of States for extended systems . . . 84
6.2.3 Free particle density of states with mass . . . 86
6.2.4 Free particle density of states without mass . . . 87
6.3 Real and reciprocal lattice . . . 88
6.3.1 Real and reciprocal lattice: One-dimensional example . . . 88
6.3.2 Real and reciprocal lattice in three dimensions . . . 89
6.4 Bloch Theorem . . . 92
6.5 Reduced zone scheme . . . 94
6.6 Bands and orbitals . . . 96
6.7 Calculating band structures in the tight-binding model . . . 98

7 From atoms to solids 103
7.1 The Atom . . . 103
7.1.1 The generalized hydrogen atom . . . 103
7.1.2 Lifting the ℓ-degeneracy: Aufbau principle . . . 104
7.2 From atoms to ions . . . 106
7.2.1 Ionization potential and electron affinity . . . 106
7.2.2 Oxidation states and electron count . . . 107

8 Instabilities 109

8.1 Jahn-Teller Instability 110
8.2 Peierl's Instability 110
8.3 Nesting 112
8.4 Other instabilities 112

9 Ionic compounds **113**
9.1 Oxidation states and charges 113
9.2 Electronegativity, hardness and softness 113
  9.2.1 The concept of electronegativity 115
  9.2.2 The concept of hardness 116
  9.2.3 Hard-Soft Acid-Base (HSAB) principle 117
9.3 Van Arkel Ketelaar triangle 117
9.4 Zintl compounds 118
9.5 Madelung constants 120
9.6 Born-Mayer Equation 120
9.7 Kapustinskii Equation 121
9.8 Born-Haber Cycle 122
9.9 Pauling bonding rules 123

10 Bond strengths **127**
10.1 Covalent bonds 127
10.2 Ionic bonds 127

11 Weak Bonds **129**
11.1 Hydrogen bonds 129
11.2 Van der Waals Bond 129

12 Crystal structures **133**
12.1 Online resources 133
12.2 Bravais lattices 133
12.3 Coordination polyhedra 134
12.4 Simple cubic structure (sic) 136
12.5 Face centered cubic structure (fcc) 136
12.6 Structure principles derived from closed packing 138
12.7 Structures derived from the fcc structure 140
  12.7.1 Zinc-blende structure (ZnS) 140
  12.7.2 Diamond structure 141
  12.7.3 Fluorite structure (CaF_2) 141
  12.7.4 Anti-fluorite structure 142
  12.7.5 Anatase structure (TiO_2) 142
  12.7.6 Bixbyite structure (Mn_2O_3) 143
  12.7.7 Rock salt (NaCl) 145
  12.7.8 Spinel structure ($MgAl_2O_4$) 145
  12.7.9 ABC-Birnessite 147
12.8 Hexagonally closed packed structure (hcp) 148
12.9 Structures derived from the hcp structure 149
  12.9.1 Graphite 149

12.9.2 NiAs structure (B8-type) 149
12.9.3 α-Corundum 149
12.9.4 Cadmium iodide 150
12.10 Body-centered cubic structure 150
12.11 Structures derived from the bcc structure 151
12.11.1 Caesium chloride structure (CsCl) 151
12.11.2 Rutile (C4-type) 152
12.12 Perovskite 152
12.12.1 YBCO 153
12.12.2 $CaCuO_2$ 154
12.13 Mixture of structural principles 155
12.13.1 Ruddlesden-Popper phases 155
12.13.2 Magnéli phases 156
12.13.3 Intermetallics 156
12.14 Additional reading 157

13 Crystal-Orbital Hamiltonian Populations 159
13.1 Definition of the density of states 159
13.2 Example: Formaldehyde 160
13.3 Example: Iron 161
13.4 Sum rules 161

14 Magnetism 165
14.1 Quick and Dirty 165
14.1.1 Driving force for a parallel (ferromagnetic) alignment of spins 165
14.1.2 Driving force for antiparallel alignment of spins 166
14.1.3 Taking things together 167
14.2 Stoner criterion 167
14.3 ???? 169
14.4 Exchange splitting 169
14.5 Zener's theory 172
14.6 Super-exchange 172
14.7 Double exchange 173

A Polyhedra in paper 181

B Justification of the Wolfsberg formula 183
B.1 Approximate orthonormalization 183
B.2 Approximate Hamiltonian with the Wolfsberg formula 184
B.3 Origin of the Wolfsberg formula 184

C Two-center bond with overlap 187
C.1 Degenerate interaction 187
C.2 Nondegenerate interaction 188
C.3 Degenerate interaction 188

D Covalent bonds to second order 189

E Atomic structures **193**
E.1 fcc 193
E.2 bcc 193
E.3 Diamond structure 193
E.4 Rocksalt NaCl structure 193
E.5 Fluorite structure 193
E.6 CsCl structure 194
E.7 Spinel 194

F Reading **197**

G Dictionary **199**

H Greek Alphabet **201**

I Philosophy of the ΦSX Series **203**

J About the Author **205**

Chapter 1

Brief review of quantum mechanics

1.1 Classical mechanics

In classical mechanics the state of a particle is described by its position $\vec{r}$ and its momentum $\vec{p}$. The total energy of a particle is described by its **Hamilton function** $H(\vec{p},\vec{r})$. In many cases the Hamilton function has the form

$$H(\vec{p},\vec{r})=\frac{p^2}{2m}+V(\vec{r}) \tag{1.1}$$

The Hamiltonian must be expressed by positions and momenta and not by positions and velocities!

The equations of motion are governed by the **Hamilton's equations**

$$\dot{p}_i=-\frac{\partial H}{\partial x_i} \qquad ; \qquad \dot{x}_i=\frac{\partial H}{\partial p_i} \tag{1.2}$$

Hamilton's equations contain the same information as Newton's equation, which can be demonstrated for the Hamilton function given above in Eq. 1.1.

$$\ddot{\vec{r}} \overset{Eq.\ 1.2}{=} \frac{d}{dt}\frac{\vec{p}}{m} \overset{Eq.\ 1.2}{=} -\frac{1}{m}\vec{\nabla}V \overset{\vec{F}=-\vec{\nabla}V}{=} \frac{1}{m}\vec{F}$$

1.2 Quantum Mechanics

Electrons in a solid must be described quantum mechanically. Quantum mechanics cannot predict positions and momenta with absolute precision. It only allows to predict probabilities for positions and momenta. The information about a quantum state is no more captured by positions and momenta but by the wave function $\Psi(\vec{r})$. The meaning of the wave function is that its absolute square is the probability that a particle is at a given position.

1.2.1 Hamilton operator and observables

Observables, i.e. observable quantities, such as energy, position, momenta etc are represented by operators that act on wave functions. The position operator is simply $\vec{r}$. The **momentum operator** is

$$\hat{p}=\frac{\hbar}{i}\vec{\nabla}$$

All observables can be represented in classical mechanics as functions of positions and momenta. The corresponding quantum mechanical operator is obtained by replacing the positions and momenta by their corresponding operators. Thus the energy operator is obtained as

$$\hat{H} = H(\hat{p}, \hat{r}) = H(\frac{\hbar}{i}\vec{\nabla}, \vec{r})$$

This operator is called **Hamilton operator**.

For the special Hamilton function given in Eq. 1.1, the Hamilton operator is

$$\hat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r})$$

1.2.2 Schrödinger equation

The dynamics of the wave function is governed by the **time-dependent Schrödinger equation**

$$i\hbar\partial_t\Psi(\vec{r}, t) = \hat{H}\Psi(\vec{r}, t)$$

If the Hamilton operator does not explicitly depend on time, we can make an ansatz for the wave functions as

$$\psi(\vec{r}, t) = \psi_E(\vec{r})\mathrm{e}^{-\frac{i}{\hbar}Et}$$

Insertion into the time-dependent Schrödinger equation yields the **time-independent Schrödinger equation**.

$$\hat{H}\Psi_i(\vec{r}) = E_i\Psi_i(\vec{r})$$

The time-independent Schrödinger equation does not have solutions for every value of the energy E. The values for which solutions exist form the **energy spectrum** of system. The spectrum can be discrete, or continuous or it may contain discrete and continuous portions.

The general solution of the time-dependent Schrödinger equation is obtained as superposition of the partial solutions of the time-independent Schrödinger equation, that is

$$\psi(\vec{r}, t) = \sum_i \psi_i \mathrm{e}^{-\frac{i}{\hbar}E_i t} c_i$$

with the complex coefficients c_i determined by the initial conditions. The initial conditions are the values of the wave function at the initial time.

1.2.3 Expectation values

Quantum mechanics does not predict the value of an observable with precision, but only the mean value of the possible results of a measurement. This mean value is called **expectation value**

$$\langle A\rangle = \frac{\int d^3r\ \psi^*(\vec{r})\hat{A}\psi(\vec{r})}{\int d^3r\ \psi^*(\vec{r})\psi(\vec{r})}$$

In order to avoid the denominator in the above expression one usually requires the wave functions to be normalized, that is

$$\int d^3r\ \psi^*(\vec{r})\psi(\vec{r}) = 1$$

The **normalization condition** reflects the requirement that the probability to find the particle anywhere in space must be equal to one.

Quantum mechanics furthermore says that a measurement of an observable A can only yield the eigenvalues of the corresponding operator. After the measurement, the system is in an eigenstate of A with the measured eigenvalue. **eigenvalue** and **eigenstate** fulfill the eigenvalue equation

$$\hat{A}\psi_A(\vec{r}) = \psi_A(\vec{r})A$$

where $\psi_A(\vec{r})$ is the eigenstate and A is the corresponding eigenvalue.

1.2.4 Bracket notation

In quantum mechanics a compact and flexible notation has been developed, namely **Dirac's bracket notation**. A wave function $\Psi(\vec{r})$ is described by a **ket** $|\psi\rangle$. The complex conjugate wave function is described by a **bra** $\langle\psi|$.

The **scalar product** between a bra $\langle\psi|$ and a ket $|\phi\rangle$ is defined as

$$\langle\psi|\phi\rangle = \int d^3r\ \psi^*(\vec{r})\phi(\vec{r}) \tag{1.3}$$

A scalar product is a complex number. Forming scalar products is the only way to extract numbers from the abstract notation. The scalar product has the general property

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$$

The **expectation value** of an operator is

$$\langle\psi|\hat{A}|\psi\rangle = \int d^3r\ \psi^*(\vec{r})\hat{A}\psi(\vec{r})$$

In order to make contact with the real-space representation described in the beginning, we introduce states

$$|\vec{r}\rangle$$

The wave function in real space is obtained from a ket by

$$\psi(\vec{r}) = \langle\vec{r}|\psi\rangle$$

The complex wave function is then

$$\psi^*(\vec{r}) = \langle\psi|\vec{r}\rangle$$

The **scalar products** have, per definition, the form

$$\langle\vec{r}|\vec{r'}\rangle = \delta(\vec{r} - \vec{r'})$$

where $\delta(\vec{r})$ is **Dirac's delta function** in three dimensions. The delta function is defined by

$$\int d^3r\ f(\vec{r})\delta(\vec{r} - \vec{r}_0) = f(\vec{r}_0)$$

which must hold for any continuous function $f(\vec{r})$.

The unit operator $\hat{1}$ is the operator that reproduces any state, to which it is applied without change. It can be expressed as

$$\hat{1} = \int d^3r\ |\vec{r}\rangle\langle\vec{r}|$$

The **position operator** has the form

$$\hat{\vec{r}} = \int d^3r\ |\vec{r}\rangle\vec{r}\langle\vec{r}|$$

The **momentum operator** has the form

$$\hat{\vec{p}} = \int d^3r\ |\vec{r}\rangle\frac{\hbar}{i}\vec{\nabla}\langle\vec{r}|$$

and the **Hamilton operator** for the Hamilton function $H = \vec{p}^2/2m + V(\vec{r})$ has the form

$$\hat{H} = H(\hat{\vec{p}}, \hat{\vec{r}}) = \int d^3r\ |\vec{r}\rangle\left(\frac{-\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r})\right)\langle\vec{r}|$$

1.3 Many non-interacting electrons and Pauli principle

If we consider many electrons, the wave function depends on all the coordinates of all electrons in the system.

Electrons have a property that plays out in quantum mechanics, namely that they are indistinguishable. Indistinguishable means that there is no experiment that allows one to distinguish one electron from the other. The result of being indistinguishable is that the wave function is either symmetric or antisymmetric under permutation of the coordinates of two particles.

$$\psi(\ldots,\vec{r}_i,\ldots,\vec{r}_j,\ldots)=\pm\psi(\ldots,\vec{r}_j,\ldots,\vec{r}_i,\ldots)$$

- Particles for which the wave function is symmetric under permutation of the particles are called **Bosons**. Bosons are particles with an integer spin. Examples are photons, phonons, gluons, gravitons, etc. Bosons are usually related to some interaction between particles.
- Particles for which the wave function is antisymmetric unter permutation are called **Fermions**. Fermions are particles with half-integer spin. Examples are electrons, protons, neutrons, quarks, etc.

A result of the antisymmetry of the wave function with respect to permutation of two particles is the **Pauli principle**: No two electrons with the same spin can be at the same place. That is, electrons with the same spin avoid each other.

It can easily be seen that the probability for two particles to be at the same place vanishes for an antisymmetric wave function. If two identical particle coordinates are changed, the wave function is the same as before. On the other hand the wave function must change its sign. Hence the wave function must vanish at this point.

Antisymmetric wave functions can be constructed using **Slater determinants**. We start out with an orthonormal set of one-particle states $\chi_i(\vec{r})$ and form

$$\Phi(\vec{r}_1,\vec{r}_2,\ldots,\vec{r}_N)=\frac{1}{\sqrt{N!}}\det\begin{vmatrix}\chi_1(\vec{r}_1) & \chi_1(\vec{r}_2) & \ldots & \chi_1(\vec{r}_N)\\ \chi_2(\vec{r}_1) & \chi_2(\vec{r}_2) & \ldots & \chi_2(\vec{r}_N)\\ \vdots & \vdots & \vdots & \vdots\\ \chi_N(\vec{r}_1) & \chi_N(\vec{r}_2) & \ldots & \chi_N(\vec{r}_N)\end{vmatrix}$$

The permutation of two particles corresponds to the permutations of two columns in the determinant. The determinant has the property that its sign changes under permutation of two columns.

The Slater determinant for two particle has the form

$$\Phi(\vec{r},\vec{r'})=\frac{1}{\sqrt{2}}\left(\chi_1(\vec{r})\chi_2(\vec{r'})-\chi_2(\vec{r})\chi_1(\vec{r'})\right)$$

We recognize immediately that the Slater determinant formed from two identical one-particle orbitals vanishes. This immediately leads to a more common form of the **Pauli principle**, namely that no two electrons with the same spin can occupy the same orbital $\chi(\vec{r})$.

Note, however, that not every antisymmetric wave function can be described by a single determinant. However, if the one-particle orbitals are complete and orthonormal, the set of all Slater determinants that can be formed from these one-particle states, is a complete basis set for the N-particle wave functions.

Hence, a general N-particle state can be expressed as a superposition of a complete set of Slater determinants.

So-far we have only considered the position of an electron. An electron is only described by its position and **spin**. The spin is the intrinsic angular momentum of the electron and it can assume two values along a certain axis, namely $S=\frac{\hbar}{2}$ and $S=-\frac{\hbar}{2}$. We refer to these electrons as spin-up and spin-down electrons.

Chapter 2

The covalent bond

2.1 Linear combination of atomic orbitals (LCAO)

The LCAO method has been proposed first by Bloch[2]. Slater and Koster[3] showed how the matrix elements can be evaluated from a small number of parameters.

Already the states of electrons in most molecules cannot be solved exactly. The problem is related to the interaction. One approximation is to assume that the electrons are not interacting. A sound theoretical justification for this approximation is given by **density functional theory (DFT)**, which maps the interacting electrons onto non-interacting electrons in an effective potential. However, density functional theory introduces additional terms in the total energy that we will not consider at this point.

We still need to determine the electronic wave functions for a complicated potential. That is we need to solve the Schrödinger equation

$$\left[\frac{-\hbar^2}{2m_e}\vec{\nabla}^2+v(\vec{r})-\epsilon_n\right]\psi_n(\vec{r})=0$$

which, in bra-ket notation has the form

$$\left[\frac{\hat{p}^2}{2m_e}+\hat{v}-\epsilon_n\right]|\psi_n\rangle=0 \tag{2.1}$$

with $\hat{v}=\int d^3r\ |\vec{r}\rangle v(\vec{r})\langle\vec{r}|$. One way to tackle this problem is to define a basis set so that we can write the one-particle wave functions $|\psi_n\rangle$ as superposition of basis functions $|\chi_{R,\ell,m,\sigma,i}\rangle$. Let us at the moment consider the wave functions of the isolated atoms as basis functions. The index R denotes a given atomic site, ℓ is the angular momentum quantum number, m the magnetic quantum number and σ is the spin quantum number. i is the principal quantum number of the atom. In the following we will use a shorthand for the indices so that $\alpha=(R,\ell,m,\sigma,i)$. Thus we represent the one-particle wave functions as

$$|\psi_n\rangle=\sum_\alpha|\chi_\alpha\rangle c_{\alpha,n} \tag{2.2}$$

We insert the ansatz Eq. 2.2 into the Schrödinger equation Eq. 2.1, and multiply the equation from the left with $\langle\chi_\beta|$. This leads to a generalized eigenvalue problem[1] for matrices

$$\sum_\alpha\left(H_{\beta,\alpha}-\epsilon_n O_{\beta,\alpha}\right)c_{\alpha,n}=0 \tag{2.3}$$

---

[1] An eigenvalue problem is called generalized if it has an overlap matrix that differs from the unity matrix

where the Hamilton matrix $H_{\alpha,\beta}$ and the overlap matrix $O_{\alpha,\beta}$ are defined by

$$\begin{aligned}H_{\alpha,\beta}&=\langle\chi_\alpha|\frac{\hat{p}^2}{2m_e}+\hat{v}|\chi_\beta\rangle\\ O_{\alpha,\beta}&=\langle\chi_\alpha|\chi_\beta\rangle\end{aligned}$$

A **generalized eigenvalue problem**, can be solved numerically using the **LAPACK** library[**?**]. The generalized eigenvalue problem yields the eigenvalues ϵ_n and the **eigenvectors** $\vec{c}_n$ with elements $c_{\alpha,n}$.

The generalized eigenvalue problem does not yet determine the norm of the eigenvectors. The normalization condition is

$$\langle\psi_n|\psi_m\rangle=\sum_{\alpha,\beta}c^*_{\alpha,n}O_{\alpha,\beta}c_{\beta,m}=\delta_{n,m}$$

2.2 From non-orthonormal to orthonormal basis functions

We consider here only one orbital on each atom as, for example, in the hydrogen molecule. We will see that the concept can be generalized to most chemical bonds. The resulting eigenvalue equation is

$$\sum_{\beta=1}^{2}H_{\alpha,\beta}c_{\beta,n}=\sum_{\beta=1}^{2}O_{\alpha,\beta}c_{\beta,n}\epsilon_n$$

where **H** is the Hamilton matrix and **O** is called the overlap matrix. The diagonal elements are approximated by the atomic eigenvalues $\bar{\epsilon}_1$ and $\bar{\epsilon}_2$ of the two atoms and the diagonal elements of the overlap matrix are unity, if we start form normalized eigenvalues. We rename $H_{12}=t$ and $O_{12}=\Delta$ so that

$$\mathbf{H}=\begin{pmatrix}\bar{\epsilon}_1 & t\\ t^* & \bar{\epsilon}_2\end{pmatrix}\qquad\text{and}\qquad\mathbf{O}=\begin{pmatrix}1 & \Delta\\ \Delta^* & 1\end{pmatrix}\tag{2.4}$$

The parameter t is called **hopping matrix element**.[2]

Approximate orthonormalization

Below, we will work with orthonormal basissets, even though the overlap is not negligible. For a non-orthonormal basis set the first step is an orthonormalization.

The Schrödinger equation $(\boldsymbol{H}-\epsilon_n\boldsymbol{O})\vec{c}_n=0$ can be rewritten by multiplication from the left with $\boldsymbol{O}^{-\frac{1}{2}}$.

$$\begin{aligned}&(\boldsymbol{H}-\epsilon_n\boldsymbol{O})\vec{c}_n=0\\ \Rightarrow\qquad&\boldsymbol{O}^{-\frac{1}{2}}(\boldsymbol{H}-\epsilon_n\boldsymbol{O})\underbrace{\boldsymbol{O}^{-\frac{1}{2}}\boldsymbol{O}^{\frac{1}{2}}}_{=\mathbf{1}}\vec{c}_n=0\\ \Rightarrow\qquad&\Big(\underbrace{\boldsymbol{O}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{O}^{-\frac{1}{2}}}_{\boldsymbol{H}'}-\epsilon_n\Big)\underbrace{\boldsymbol{O}^{\frac{1}{2}}\vec{c}_n}_{\vec{c}'}=0\end{aligned}$$

A function of an operator, such as the inverse square root, is defined via its Taylor expansion. The Taylor expansion of $f(x)=(1+x)^{-\frac{1}{2}}$ is $f(x)=1-\frac{1}{2}x+O(x^2)$. Thus the inverse square roort of the overlap operator is

---

[2] The name results from the picture of an electron hopping from site to site. This transport is larger if the hopping matrix element is larger. It is often used in the context of orthonormal orbitals. The picture of and electron performing discrete hops from site to site is, however, misleading.

$$\begin{pmatrix} 1 & \Delta \\ \Delta^* & 1 \end{pmatrix}^{-\frac{1}{2}} = \begin{pmatrix} 1 & -\frac{1}{2}\Delta \\ -\frac{1}{2}\Delta^* & 1 \end{pmatrix} + O(|\Delta|)^2$$

This gives us the transformed Hamiltonian in the form

$$\boldsymbol{H'} = \begin{pmatrix} \bar{\epsilon}_1 - \mathrm{Re}[\Delta^* t] & t - \Delta\frac{\bar{\epsilon}_1+\bar{\epsilon}_2}{2} \\ t^* - \Delta^*\frac{\bar{\epsilon}_1+\bar{\epsilon}_2}{2} & \bar{\epsilon}_2 - \mathrm{Re}[\Delta^* t] \end{pmatrix} + O(|\Delta|^2) \tag{2.5}$$

Normally the hopping matrix element t is negative and the overlap matrix element Δ is positive. We then observe that the diagonal elements are shifted upwards in energy. This effect is called **Pauli repulsion**: If two atoms come close, their atomic orbitals overlap. When orthogonality is restored, the energy levels shift up. A simple argument goes as follows: As two atoms overlap, the electrons repel each other due to the Pauli principle. The atoms become effectively compressed. As a consequence, via Heisenberg's uncertainty principle[3], the kinetic energy is increased and energy is shifted up.

Wolfsberg-Helmholtz Formula

We may use a simple empirical relation for the hopping matrix elements, namely the **Wolfsberg-Helmholtz formula**[4], which says

$$t = k\frac{\bar{\epsilon}_1 + \bar{\epsilon}_2}{2}\Delta \tag{2.6}$$

where $k \approx 1.75$ is an empirical constant. Values for k depend somewhat on the type of the bond and vary between 1.6 and 2.0. For a motivation of the Wolfsberg formula see App. B on p. 183.

One can now orthonormalize the orbitals. Using the Wolfsberg-Helmholtz formula Eq. 2.6 we obtain a Hamiltonian Eq. 2.5 to first order in Δ, which has the form

$$\mathbf{H}' = \begin{pmatrix} \bar{\epsilon}_1 & (k-1)\frac{\bar{\epsilon}_1+\bar{\epsilon}_2}{2}\Delta \\ (k-1)\frac{\bar{\epsilon}_1+\bar{\epsilon}_2}{2}\Delta^* & \bar{\epsilon}_2 \end{pmatrix} \quad \text{and} \quad \mathbf{O}' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The derivation is given in App. B.

The main difference from a general Hamilton matrix given in Eq. 2.4 is the renormalization of the hopping matrix element from t to $t' = (k-1)\frac{\bar{\epsilon}_1+\bar{\epsilon}_2}{2}\Delta$.

The change of the diagonal elements amounts to a upward shift of both energy levels by the same amount. The shift is (usually) upward, because the atomic energy levels are (usually) negative.

Complete neglect of overlap (CNO)

Since we want to approach the problem in small steps, we go back to Eq. 2.4 and ignore the off-diagonal elements of the overlap matrix. Thus, in the following we consider instead

$$\mathbf{H} = \begin{pmatrix} \bar{\epsilon}_1 & t \\ t^* & \bar{\epsilon}_2 \end{pmatrix} \quad \text{and} \quad \mathbf{O} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.7}$$

For our discussion it is important that $t < 0$, which is usually fulfilled. This approximation is called **Complete Neglect of Overlap** (CNO).

---

[3] Heisenberg's uncertainty principle says that the variation of position and momentum is equal or larger than $\frac{1}{2}\hbar$, that is $\Delta x \Delta p \geq \frac{1}{2}\hbar$. This suggests that the momentum of a confined particle should be larger than about $p > \frac{\hbar}{2\Delta x}$. Hence the energy would be about $E_{kin} > \frac{\hbar^2}{m(2\Delta x)^2}$. This says that the kinetic energy rises with confinement.

2.3 The two-center bond

. We consider now the simple case of a system with two orbitals. The typical example is a hydrogen molecule. However the findings will be more general and will be applicable in most cases where two atoms form a bond.

The model Hamiltonian, we investigate, is

$$\mathbf{H}=\begin{pmatrix}\bar\epsilon_1 & t\\ t^* & \bar\epsilon_2\end{pmatrix}\qquad\text{and}\qquad \mathbf{O}=\mathbf{1}=\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

Eigenvalues

We diagonalize the Hamiltonian by finding the zeros of the determinant of $\mathbf{H}-\epsilon\mathbf{1}$.[4] The zero's of the characteristic polynomial determine the eigenvalues

$$0=(\bar\epsilon_1-\epsilon)(\bar\epsilon_2-\epsilon)-|t|^2$$

EIGENVALUES OF THE TWO-CENTER BOND

$$\epsilon_\pm=\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\pm\sqrt{\left(\frac{\bar\epsilon_1-\bar\epsilon_2}{2}\right)^2+|t|^2}\tag{2.8}$$

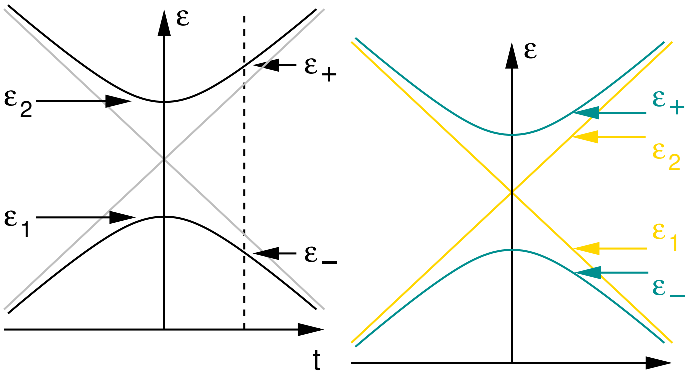


Fig. 2.1: Left: Energy levels of the two-center bond as function of the hopping parameter t. Right: energy levels as functiuon of some parameter, which tunes the diagonal elements of the Hamiltonian at a fixed hopping parameter t.

The result of Eq. 2.8 is shown in Fig. 2.1. If the hopping parameter t vanishes, the eigenvalues are, naturally, just the diagonal elements of the Hamiltonian, the "atomic energy levels" $\bar\epsilon_1$ and $\bar\epsilon_2$. With increasing hopping parameter the splitting of the energy levels grows. It never becomes smaller!

- for a large hopping parameter t, i.e. for $|t|>>|\bar\epsilon_2-\bar\epsilon_1|$, we obtain approximately

  $$\epsilon_\pm\approx\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\pm|t|\tag{2.9}$$

  If the hopping parameter becomes much larger than the initial energy level splitting $\bar\epsilon_2-\bar\epsilon_1$, the energy levels deviate approximately linearly with t from the mean value $(\bar\epsilon_1+\bar\epsilon_2)/2$.

---

[4] The condition that the determinant vanishes, that is $\det[\mathrm{H}-\epsilon\mathrm{O}]=0$ determines the eigenvalues of the system. We need to determine the zeroes of a polynomial of the energy. This polynomial is called the **characteristic polynomial**

- for a small hopping parameter t, i.e. for $|t| << |\epsilon_2 - \epsilon_1|$, we obtain approximately

$$\epsilon_- \approx \bar{\epsilon}_1 - \frac{|t|^2}{|\epsilon_2 - \epsilon_1|} \tag{2.10}$$

$$\epsilon_+ \approx \bar{\epsilon}_2 + \frac{|t|^2}{|\epsilon_2 - \epsilon_1|} \tag{2.11}$$

  For small t the energy levels $\epsilon_\pm$ deviate approximately quadratically with t from the "atomic energy levels" $\bar{\epsilon}_{1/2}$. The level shift is larger if the energy levels lie close initially.

Diagonalization conserves the trace

The mean value of the eigenvalues remains always the same, if basisset is orthonormal. This is a consequence of the fact that the trace of a matrix is invariant under a unitary transformation.

The normalized eigenvectors $\vec{c}_n$ of a hermitean matrix $\boldsymbol{H}$ form a unitary matrix $\boldsymbol{U}$ with $U_{\alpha,n} = c_{\alpha,n}$. Thus, the eigenvalue equation has the form

$$\boldsymbol{HU} = \boldsymbol{Uh}$$

where $\boldsymbol{h}$ is a diagonal matrix with the eigenvalues ϵ_n on the main diagonal and $\boldsymbol{U}$ is unitary, that is $\boldsymbol{UU}^\dagger = \mathbf{1}$. When we exploit that the trace of a product of matrices or operators is invariant under cyclic permutations of the individual terms, we can show

$$\sum_\alpha H_{\alpha,\alpha} = \mathrm{Tr}[\boldsymbol{H}] = \mathrm{Tr}[\boldsymbol{HUU}^\dagger] = \mathrm{Tr}[\boldsymbol{UhU}^\dagger] = \mathrm{Tr}[\boldsymbol{U}^\dagger\boldsymbol{Uh}] = \mathrm{Tr}[\boldsymbol{h}] = \sum_n \epsilon_n$$

which proves that the sum of eigenvalues is identical to the sum of diagonal elements of the Hamiltonian. Hence, the sum of energy levels is equal to the sum of "atomic" energy levels.

This statement says that the stabilizing effect of a bond is exactly canceled by the destabilizing effect of the corresponding antibond. It is, however, only true for an orthonormal basis set. If the overlap matrix is not unity, there is the Pauli repulsion shifting the orbitals upward.

Eigenvectors

The two eigenvectors $\vec{c}_n$ with $n \in \{+,-\}$ are obtained from $(\boldsymbol{H} - \epsilon_n\mathbf{1})\vec{c}_n = 0$, that is alternatively from the equation

$$(\bar{\epsilon}_1 - \epsilon_\pm)c_{1,\pm} + tc_{2,\pm} = 0 \tag{2.12}$$

or from

$$t^* c_{1,\pm} + (\bar{\epsilon}_2 - \epsilon_\pm)c_{2,\pm} = 0 \tag{2.13}$$

Both equations lead to the same result.

In our two-dimensional case, we can solve the equations simply by looking for an orthogonal vector to the coefficients[5] and to normalize it. Because the results will have a more transparent form, we choose the second equation Eq. 2.13 for the lower, bonding eigenstate

$$c_{1,-} = \frac{\bar{\epsilon}_2 - \epsilon_-}{\sqrt{(\bar{\epsilon}_2 - \epsilon_-)^2 + |t|^2}} \qquad \text{and} \qquad c_{2,-} = \frac{-t^*}{\sqrt{(\bar{\epsilon}_2 - \epsilon_-)^2 + |t|^2}} \tag{2.14}$$

and the first equation Eq. 2.12 for the higher, antibonding state

$$c_{1,+} = \frac{t}{\sqrt{(\bar{\epsilon}_1 - \epsilon_+)^2 + |t|^2}} \qquad \text{and} \qquad c_{2,+} = \frac{-(\bar{\epsilon}_1 - \epsilon_+)}{\sqrt{(\bar{\epsilon}_1 - \epsilon_+)^2 + |t|^2}} \tag{2.15}$$

---

[5] A complex 2-dimensional vector $\vec{b}$ orthogonal to a vector $\vec{a}$ can be found simply as $b_1 = a_2$ and $b_2 = -a_1$

Both results can be combined into

EIGENVECTORS OF THE TWO-CENTER BOND

$$\vec{c}_- = \begin{pmatrix} 1 \\ \frac{-t^*}{\bar{\epsilon}_2-\epsilon_-} \end{pmatrix} \left(1+\frac{|t|^2}{(\bar{\epsilon}_2-\epsilon_-)^2}\right)^{-\frac{1}{2}} \quad \text{and} \quad \vec{c}_+ = \begin{pmatrix} \frac{t}{\epsilon_+-\bar{\epsilon}_1} \\ 1 \end{pmatrix} \left(1+\frac{|t|^2}{(\epsilon_+-\bar{\epsilon}_1)^2}\right)^{-\frac{1}{2}} \tag{2.16}$$

The eigenvalues $\epsilon_\pm$ are given in Eq. 2.8.

In order to to make the qualitatively relevant results more evident, we distinguish the degenerate limit, i.e. $(\epsilon_2 - \epsilon_1) << |t|$ from the non-degenerate case $(\epsilon_2 - \epsilon_1) >> |t|$.

Degenerate case

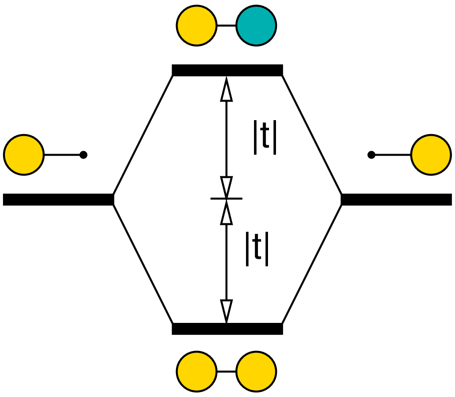
In the degenerate limit the two "atomic levels" are identical, i.e. $\bar{\epsilon}_1 = \bar{\epsilon}_2 =: \bar{\epsilon}$. The degenerate case describes the bonding of two symmetric orbitals, such as the orbitals of a hydrogen molecule.

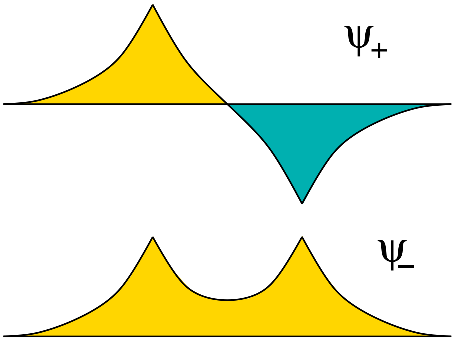
From Eq. 2.8 we can directly determine the energy eigenvalues as

EIGENVALUES OF THE DEGENERATE TWO-CENTER BOND

$$\epsilon_\pm = \bar{\epsilon} \pm |t| \tag{2.17}$$

The lower wave function with energy ϵ_- is is the **bonding state** and the upper wave function with energy ϵ_+ is called the **anti-bonding state**.





Let us consider the binding between two atoms with one electron each. Before the bond is formed, the atoms are far apart and the hopping matrix element t vanishes. Once the bond has formed, both electrons can move into the lower, bonding orbital. The energy gained is $2|t|$. If there are two electrons in each orbital or if there are no electrons the sum of occupied energy eigenvalues remains identical. Occupying the anti-bonding orbital (at $\bar{\epsilon} + |t|$) costs energy.

We can look up the eigenvectors from Eq. 2.16, but they are easily obtained directly[6]:

$$\begin{pmatrix} \bar{\epsilon}-\epsilon_\pm & t \\ t^* & \bar{\epsilon}-\epsilon_\pm \end{pmatrix} \begin{pmatrix} c_{1,\pm} \\ c_{2,\pm} \end{pmatrix} = \begin{pmatrix} \mp|t| & t \\ t^* & \mp|t| \end{pmatrix} \begin{pmatrix} c_{1,\pm} \\ c_{2,\pm} \end{pmatrix} = 0 \qquad \Rightarrow \qquad c_{2,\pm} = \pm\frac{t}{|t|}c_{1,\pm}$$

---

[6] Personally, I prefer the direct calculation, because the steps are easier to memorize than a formula

If the hopping parameter t is real and negative, we obtain the eigenstates

$$
\begin{aligned}
|\Psi_-\rangle &= (|\chi_1\rangle + |\chi_2\rangle)\frac{1}{\sqrt{2}}\\
|\Psi_+\rangle &= (|\chi_1\rangle - |\chi_2\rangle)\frac{1}{\sqrt{2}}
\end{aligned}
$$

The state $|\Psi_-\rangle$ with lower energy is the bonding state and the state $|\Psi_+\rangle$ is the anti-bonding state.

The anti-bonding wave function has one node-plane, while the bonding wave function has none. A **node plane** is that surface, where the wave function changes its sign.

The bonding wave function is stabilized, because the electron can spread over two sites: According to Heisenberg's uncertainty principle[7], this spreading out leads to a lower kinetic energy.[8]

The fact that the anti-bonding wave function has a node, indicates that it has a higher kinetic energy than the bonding state. Kinetic energy can be looked upon as a measure for the mean square curvature of the wave function. This curvature becomes larger nodes are introduced.

We can now estimate the bond energy as function of the number of electrons. For each electron in the bonding orbital we gain an energy t and for each electron in the anti-bonding orbital we loose an energy t. Thus every electron in an anti-bonding orbital cancels the stabilization of an electron in the bonding orbital.

Non-degenerate case

Let us now consider the non-degenerate state. We use Eq. 2.8 with the choice $\epsilon_1 < \epsilon_2$ and truncate the Taylor expansion of the result in t after the second order. $|t|$.

$$
\epsilon_\pm = \frac{\bar\epsilon_1+\bar\epsilon_2}{2} \pm \sqrt{\left(\frac{\bar\epsilon_1-\bar\epsilon_2}{2}\right)^2+|t|^2} = \frac{\bar\epsilon_1+\bar\epsilon_2}{2} \pm \left[\frac{\bar\epsilon_2-\bar\epsilon_1}{2} + \frac{|t|^2}{\bar\epsilon_2-\bar\epsilon_1} + O(|t|^4)\right]
$$

APPROXIMATE EIGENVALUES OF THE NON-DEGENERATE TWO-CENTER BOND

$$
\epsilon_- = \bar\epsilon_1 - \frac{|t|^2}{\bar\epsilon_2-\bar\epsilon_1} + O(|t|^4) \tag{2.18}
$$

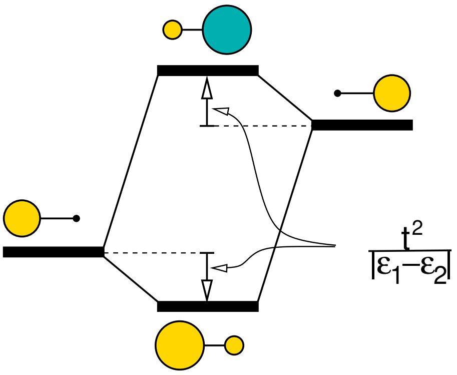
$$
\epsilon_+ = \bar\epsilon_2 + \frac{|t|^2}{\bar\epsilon_2-\bar\epsilon_1} + O(|t|^4) \tag{2.19}
$$

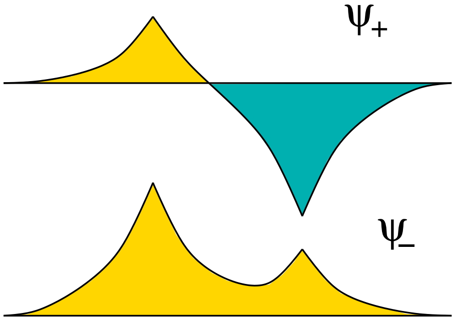
If we start from one electron in each orbital, the energy gain consists of two parts. First we gain an amount $\bar\epsilon_2 - \bar\epsilon_1$ by transferring the electron from the upper orbital at $\bar\epsilon_2$ to the lower orbital at $\bar\epsilon_1$. This is the **ionic contribution**, because a cation and an anion are formed. Secondly the lower orbital is lowered through **hybridization** with the higher orbital and we gain $\frac{2|t|^2}{\bar\epsilon_2-\bar\epsilon_1}$ for the electron pair. This covalent contribution becomes smaller the larger the initial energy separation. Thus if the ionic contribution is large, the covalent contribution is usually small.

---

[7] Heisenberg's uncertainty principle says in a specialized version, that $\Delta x \Delta p \geq \frac{\hbar}{2}$.

[8] The reader may argue that the antibonding orbital is spread out over an even larger volume but has a higher energy. Extending the argument based on Heisenberg's uncertainty principle to antibonding states can be done, but it appears a bit artificial: Consider a particle in a box. In that case the atomic case corresponds to a box only half as large. Thus the volume for the electron becomes larger in the molecule, that is the larger box. The node-plane restricts the "effectively accessible volume". Thus the antibonding state has higher energy. In a real molecule the electron can avoid some of this penalty by moving into a region with higher potential, but this can be considered a secondary effect.





The eigenvectors can be obtained from Eq. 2.16, but approximate eigenstates are easily obtained directly by inserting the approximate eigenvalues into

$$\begin{pmatrix} \bar{\epsilon}_1-\epsilon_\pm & t \\ t^* & \bar{\epsilon}_2-\epsilon_\pm \end{pmatrix}\begin{pmatrix} c_{1,\pm} \\ c_{2,\pm} \end{pmatrix}=0$$

We use the upper row for the lower eigenvalue ϵ_- and insert the approximate value from Eq. 2.19.

$$\frac{|t|^2}{\bar{\epsilon}_2-\bar{\epsilon}_1}c_1+tc_2=0 \qquad \Rightarrow \qquad c_2=\frac{-t^*}{\bar{\epsilon}_2-\bar{\epsilon}_1}c_1$$

Analogously, we obtain the eigenstate for the upper eigenvalue from the second row of the matrix equation. The normalization only enters in second order of $|t|$, and will be neglected. The result can be compared to the Taylor expansion of Eq. 2.16.

Let us now consider the eigenstates

$$\begin{aligned} |\Psi_-\rangle &\approx |\chi_1\rangle+|\chi_2\rangle\frac{-t^*}{\bar{\epsilon}_2-\bar{\epsilon}_1} \\ |\Psi_+\rangle &\approx |\chi_2\rangle-|\chi_1\rangle\frac{-t}{\bar{\epsilon}_2-\bar{\epsilon}_1} \end{aligned}$$

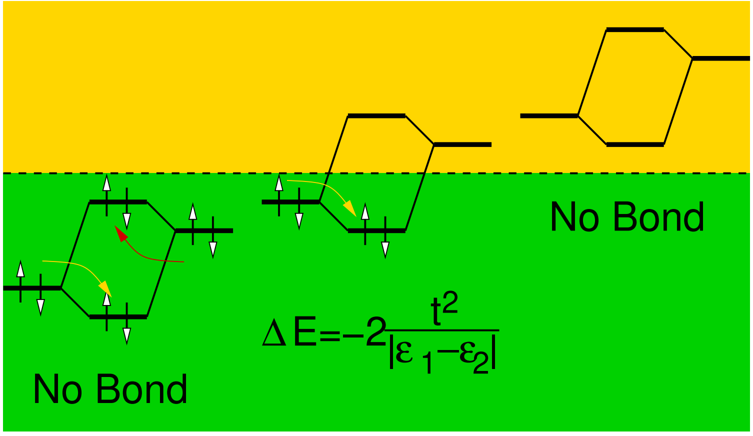
The bonding state is mostly localized on the atom with lower energy, but both orbitals contribute with the same sign. The anti-bonding state is localized mostly on the atom with higher energy and, as in the degenerate case there is a node plane between the atoms.

2.4 Bonds and occupations

In section 2.3 on p 19 we have shown, that the destabilization of the anti-bonding orbital is identical to the stabilization of the bonding orbital. The argument, which is only valid for orthonormal basissets, was traced to the invariance of the trace of a matrix under a unitary transformation. A diagonalization of a Hamiltonian, is a unitary transformation from the original basis set to the eigenstates of the Hamiltonian. Thus the sum of the eigenvalues is equal to the sum of the Hamilton expectation values of all basis functions.

Thus, we gain energy only if the orbitals are partially occupied. This statement has important consequences:

- Orbitals that lie far from the **Fermi-level** do not contribute to bonding. This explains why the role of core states to binding can be ignored. The interaction of core states from different atoms is negligible because bonding and anti-bonding orbitals are occupied.



- Molecules with unpaired electrons are very reactive because they have an orbital close to the Fermi level, that can interact with both, filled and empty, orbitals. Such molecules are called **radicals**. An example for a radical is the hydroxyl radical OH. Because radicals are so reactive they are very rare. Nevertheless they often play an important role as intermediate in chain reactions. Thus most orbitals exhibit paired electrons

- Molecules with an empty orbital just above the Fermi level are called an **electrophile**. Molecules with a filled orbital just below the Fermi level are called a **nucleophile**. An electrophile can only form a bond with a radical or a nucleophile. If two electrophile come together no bond can be formed because both orbitals are empty. If two nucleophiles come together, they cannot form a bond because a bond would result in occupied bonding as well as anti-bonding states, so that the net stabilization would vanish. The name electrophile indicates that the molecule "likes electrons", namely those of the nucleophile.

- Molecules can exhibit electrophilic and nucleophilic behavior at the same time. The electrophilic and nucleophilic orbitals are called **frontier orbitals**[5]. These are the orbitals that control the reactivity of a molecule. Usually these frontier orbitals are the **highest occupied molecular orbital (HOMO)** and the **lowest unoccupied molecular orbital (LUMO)**.

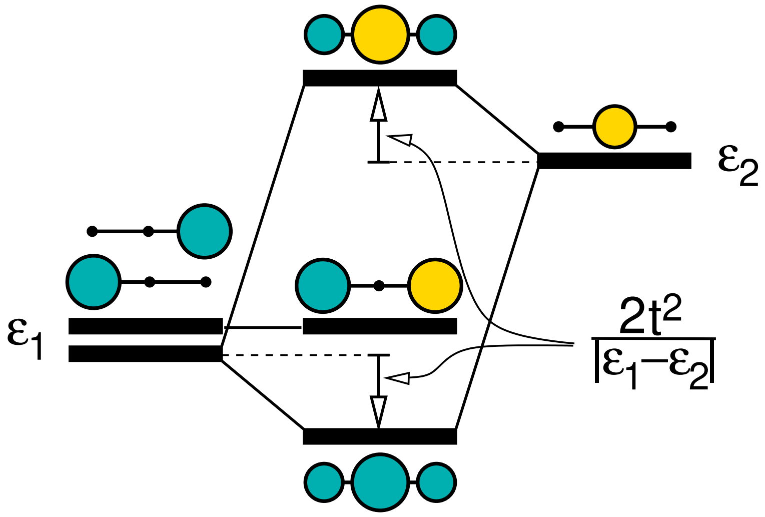
- Molecules with a large band gap between occupied and unoccupied states tend to be very stable, while molecules with orbitals near the Fermi level tend to be reactive.

2.5 Three-center bond

While most bonds can be characterized as two-center bonds, one often encounters bonds between three centers. Let us consider a central atom with two ligands on either side.

The Hamiltonian has the form

$$\mathbf{H}=\begin{pmatrix}\epsilon_1 & t & 0\\ t^* & \epsilon_2 & t\\ 0 & t^* & \epsilon_1\end{pmatrix}$$



The characteristic equation det$[\mathbf{H}-\epsilon\mathbf{1}]=0$ is

$$\begin{aligned}(\epsilon_1-\epsilon)\left[(\epsilon_2-\epsilon)(\epsilon_1-\epsilon)-|t|^2\right]-t^2(\epsilon_1-\epsilon)&=0\\(\epsilon_1-\epsilon)\left[(\epsilon_2-\epsilon)(\epsilon_1-\epsilon)-2|t|^2\right]&=0\end{aligned}$$

$$\epsilon=\begin{cases}\frac{\epsilon_1+\epsilon_2}{2}-\frac{|\epsilon_1-\epsilon_2|}{2}\sqrt{1+\left(\frac{2\sqrt{2}t}{\epsilon_1-\epsilon_2}\right)^2}\approx\epsilon_1-\frac{2t^2}{|\epsilon_1-\epsilon_2|}\\\epsilon_1\\\frac{\epsilon_1+\epsilon_2}{2}-\frac{|\epsilon_1-\epsilon_2|}{2}\sqrt{1+\left(\frac{2\sqrt{2}t}{\epsilon_1-\epsilon_2}\right)^2}\approx\epsilon_2+\frac{2t^2}{|\epsilon_1-\epsilon_2|}\end{cases}$$

We can see that there are three states: a bonding state, a non-bonding state and an anti-bonding state. Due to symmetry, the non-bonding state cannot interact with the orbital on the central atom.

Characteristic for a three center bond is that it maintains its full bond strength for two, three and four electrons, because the electrons enter into the non-bonding orbital, which does not contribute to the bond-strength.

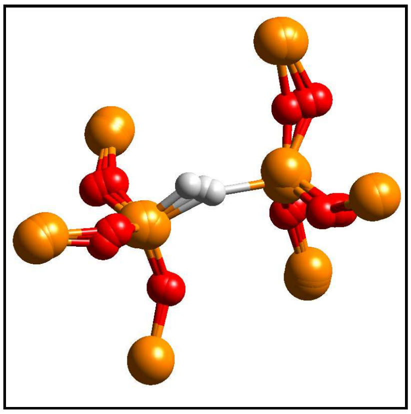


Fig. 2.2: The hydrogen complex with an oxygen vacancy in silica is responsible for stress-induced leakage currents in transistors. Because the electrons tunnel through a non-bonding orbital of the three-center bond Si-H-Si, electrons are not trapped. Red balls are oxygen atoms, yellow balls are silicon atoms and the white ball is a hydrogen atom. Structures for the three charge states +/0/- are superimposed. Only the relevant atoms out of the infinite crystal are shown. From [6, 7].

2.6 Chains and rings

Let us now extend our description from a three-center bond to a chain and ring structures of many atoms.

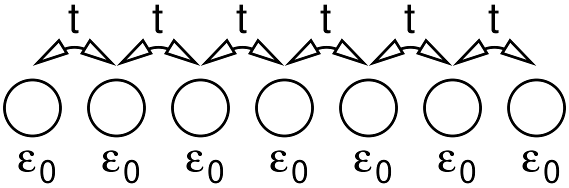
A chain is a model for a polymer, such as polyacetylene[9] It is also a model for a finite cluster. It is also a model for a state in a semiconductor heterojunction, which is within the conduction of one material, but in the band gap of the two neighboring materials.

A ring is a model for example for aromatic molecules such as benzene.

Chains and rings will later be important for the description of solids. For the description we will use **periodic boundary conditions,** two describe an infinite crystal. In one dimension, this construction corresponds directly to a ring structure. Similarly, the chain is a one-dimensional model for a solid with surfaces.

2.6.1 Chains

We assume that all atoms are identical.



Here and in the following we assume that the hopping parameter t is a real number

The Hamiltonian of the linear chain has the following form

$$\mathbf{H}=\begin{pmatrix}\epsilon_0 & t & 0 & 0 & \cdots\\ t & \epsilon_0 & t & 0 & \cdots\\ 0 & t & \epsilon_0 & t & \cdots\\ 0 & 0 & t & \epsilon_0 & \cdots\\ \vdots & \vdots & & & \end{pmatrix}\qquad \text{with } \mathrm{Im}[t]=0$$

The Schrödinger equation $(\mathbf{H}-\epsilon\mathbf{1})\vec{c}=0$ can be written line-by-line as

$$(\epsilon_0-\epsilon)c_1+tc_2=0 \tag{2.20}$$
$$tc_{i-1}+(\epsilon_0-\epsilon)c_i+tc_{i+1}=0 \qquad \text{for} \qquad 2<i<N-1 \tag{2.21}$$
$$tc_{N-1}+(\epsilon_0-\epsilon)c_N=0 \tag{2.22}$$

---

[9] Polyacetylene $(CH)_n$ dimerizes. The chain may describe the *pi* orbitals for the undimerized polyacetylene. For the dimerized system the antibonding orbitals of a dimer may play the role of one orbital in our model

Let us make a little detour, before we continue: Eq. 2.21 has similarities with a differential equation. Introducing a small spacing Δ, it can be written in the form

$$t\Delta^2 \underbrace{\frac{c_{i-1} - 2c_i + c_{i+1}}{\Delta^2}}_{\approx \partial_x^2 f(x)} + (\epsilon_0 + 2t)c_i - \epsilon c_i = 0$$

If we consider a function $f(x)$ with values $f(j\Delta) = c_j$, we can look at the above equation as a discretized version of the following differential equation for $f(x)$

$$\left[t\Delta^2 \partial_x^2 + (\epsilon_0 + 2t) - \epsilon\right] f(x) = 0$$

This equation is analogous to the one-dimensional Schrödinger equation for a constant potential. The solutions of this problem are plane waves, which we can exploit to find a solution for the original problem.

Dispersion relation

After this intermezzo let us continue with the equation for the orbital coefficients. The translational symmetry suggests that we use an exponential ansatz for Eq. 2.21.

$$c_j = \mathrm{e}^{ik\Delta j} \tag{2.23}$$

We have inserted here the spacing Δ between the sites, so that the parameter k has the usual meaning and units of a wave vector.

Insertion of this ansatz into Eq. 2.21 yields

$$\begin{aligned} t\mathrm{e}^{-ik\Delta} + (\epsilon_0 - \epsilon) + t\mathrm{e}^{ik\Delta} &\overset{Eqs.\ 2.21,2.23}{=} 0 \\ \Rightarrow 2t\cos(k\Delta) + (\epsilon_0 - \epsilon) &= 0 \\ \Rightarrow \epsilon(k) &= \epsilon_0 + 2t\cos(k\Delta) \end{aligned} \tag{2.24}$$

This is the **dispersion relation** $\epsilon(k)$ for the linear chain, shown in Fig. 2.3.

For small molecules one can investigate the energy levels individually. However, for complex molecules or crystals, the energy levels are positioned so close in energy that such a representation is no more useful. Therefore one chooses a different representation, namely the **Density of States**. As the name says the density of states is the density of energy levels as function of energy. For s molecule with discrete energies, the density of states would be a sum of δ-functions, one for each energy level. For our chain, we will see below that the allowed k-values are equispaced on the k-axis. Using the dispersion relation $\epsilon(k)$, we can determine the spacing of energy levels on the energy axis.

$$\Delta\epsilon = \underbrace{\frac{d\epsilon}{dk}}_{v_g} \underbrace{\Delta k}_{\frac{\pi}{(N+1)a}}$$

which gives us the density of states as

$$D(\epsilon) = \frac{1}{\Delta\epsilon} = \left(\frac{d\epsilon}{dk}\right)^{-1} \frac{1}{\Delta k}$$

Thus the density of states of a one-dimensional system is proportional to the inverse slope of the dispersion relation. The shape of the density of states shown in Fig. 2.3 is characteristic for a one-dimensional problem. For two or three dimensional problems the density of states would not diverge at the band edges, but start with a step in two dimensions or like the square root of the energy relative to the band edge in three dimensions.

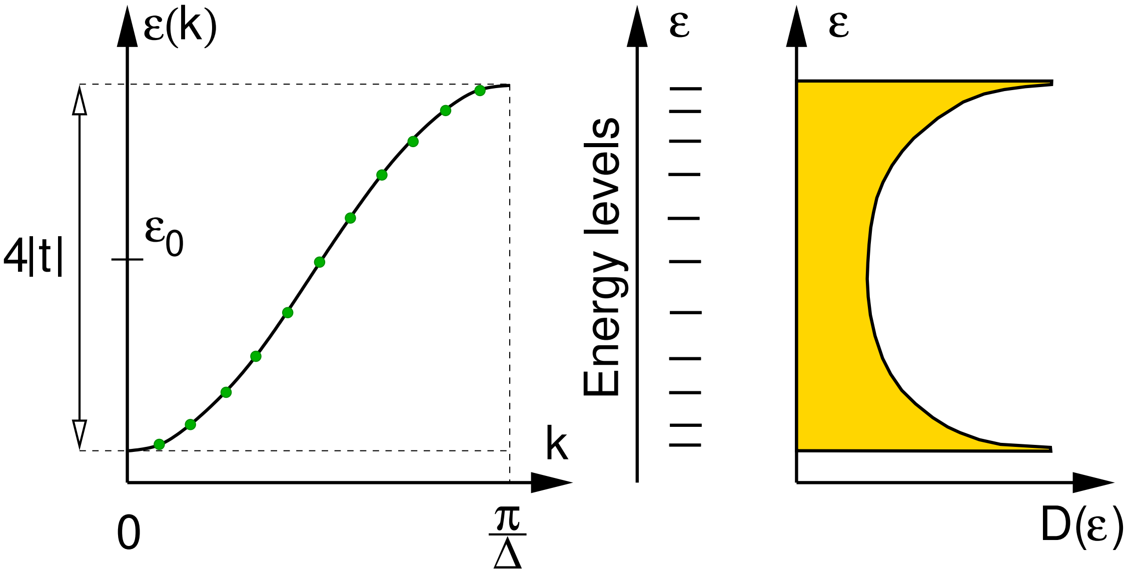


Fig. 2.3: Dispersion relation of the linear chain. The green points correspond to the energy levels of a chain with 11 atoms. In the middle figure we show the energy levels, and on the right the schematic density of states $D(\epsilon)$ for the infinite chain.

Boundary conditions

We still need to enforce the boundary conditions Eqs. 2.20 and 2.22. We can bring the boundary conditions into the form of the central condition Eq. 2.21.

$$(\epsilon_0-\epsilon)c_1+tc_2 \overset{\text{Eq. 2.20}}{=} 0 \Rightarrow \begin{cases} tc_{i-1}+(\epsilon_0-\epsilon)c_i+tc_{i+1}=0 & \text{for } i=1 \\ \text{and} \qquad c_0=0 \end{cases} \tag{2.25}$$

$$tc_{N-1}+(\epsilon_0-\epsilon)c_N \overset{\text{Eq. 2.22}}{=} 0 \Rightarrow \begin{cases} tc_{i-1}+(\epsilon_0-\epsilon)c_i+tc_{i+1}=0 & \text{for } i=N \\ \text{and} \qquad c_N=0 \end{cases} \tag{2.26}$$

Thus, the middle equation is now extended also to the end points $i=1$ and $i=N$, which were not covered by Eq. 2.21. Two additional sites with $i=0$ and $i=N+1$ have been introduced in addition.

The boundary conditions, on the other hand obtain the particularly simple form

$$c_0=0 \qquad \text{and} \qquad c_{N+1}=0 \tag{2.27}$$

I like this form, because it is now analogous to the well known particle-in-a-box problem. The only difference is the disrete nature of the problem, namely the discrete set of orbitals as opposed to a continuous range of a spatial coordinate.

For a given energy we obtain two linear independent solutions, which we superimpose to form a general ansatz.

$$c_j=Ae^{ik\Delta j}+Be^{-ik\Delta j} \tag{2.28}$$

The boundary conditions, now in the form $c_0=c_{N+1}=0$, determine the allowed values for k and $\frac{B}{A}$.

- The first boundary condition, $c_0=0$

$$0=c_0 \overset{\text{Eq. 2.28}}{=} A+B \qquad \Rightarrow \qquad B=-A$$

restricts the Ansatz Eq. 2.28 for the wave function to pure sine functions, i.e.

$$c_j = 2iA\sin(k\Delta j) \tag{2.29}$$

- the second boundary condition, $c_{N+1} = 0$

  $$0 = c_{N+1} = 2iA\sin(k\Delta(N+1))$$

  restricts the wave vectors k to the discrete values

  $$k_n = \frac{\pi}{\Delta(N+1)}n \qquad \text{for integer } n$$

Exclude trivial quantum numbers

Many solutions with different quantum number n are actually identical. This can be seen as follows: Consider the orbital coefficients in Eq. 2.30.

- periodicity of the sine function

  $$\begin{aligned}\sin(k_n\Delta j) = \sin(k_n\Delta j + 2\pi j) = \sin\left(\left[k_n + \frac{2\pi}{\Delta}\right]\Delta j\right) &\overset{\text{Eq. 2.32}}{=} \sin\left(\left[\frac{\pi}{\Delta(N+1)}n + \frac{\pi}{\Delta(N+1)}2(N+1)\right]\Delta j\right)\\ = \sin\left(\frac{\pi}{\Delta(N+1)}\Big[n + 2(N+1)\Big]\right) &= \sin\left(k_{n+2(N+1)}\Delta j\right)\end{aligned}$$

  This shows that the adding $2(N+1)$ to a quantum number n does not produce a new state. Thus we can limit the quantum numbers N to the interval $-N, \ldots, N+1$

- antisymmetry of the sine function

  $$\sin(k_n\Delta j) = -\sin(-k_n\Delta j) \overset{\text{Eq. 2.32}}{=} -\sin(k_{-n}\Delta j)$$

  Thus we can exclude all negative quantum numbers, that is all from $-N$ to -1.

We are left with quantum numbers $0, 1, \ldots, N+1$.

Two of the $N+2$ quantum numbers produce a wave function that is zero, namely $n = 0$ and $n = N+1$. For this reason we limit the quantum numbers to the range from 1 to N.

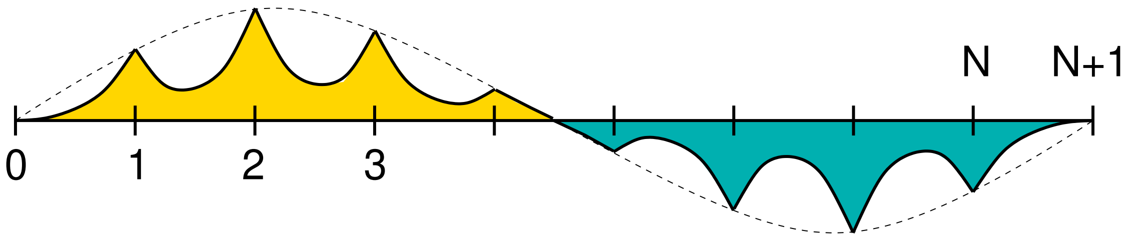
STATES OF THE LINEAR CHAIN

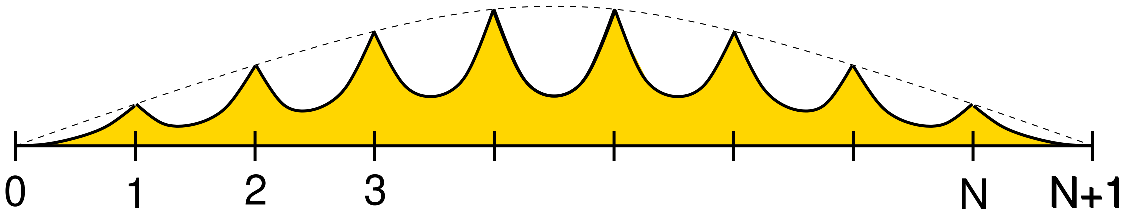
Thus the wave functions and energies are

$$|\psi_n\rangle = \sum_j |\chi_j\rangle \sin(k_n\Delta j)\sqrt{\frac{2}{N+1}} \tag{2.30}$$

$$\epsilon_n = \epsilon_0 + 2t\cos(k_n\Delta) \tag{2.31}$$

$$k_n = \frac{\pi}{\Delta(N+1)}n \qquad \text{with} \qquad n = 1, 2, \ldots, N \tag{2.32}$$

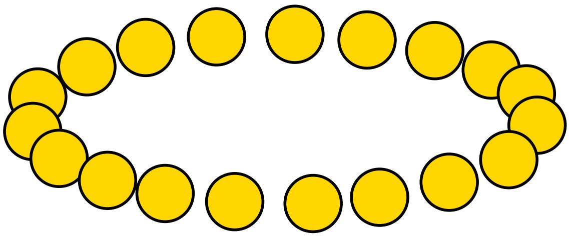




Note that the envelope of the wave function corresponds directly to the states of a particle in a box. This is a direct consequence of the analogy of Eq. 2.21 with a particle in a constant potential discussed before.

Note that the three center bond is a simple example for a linear chain. It is instructive to compare the general results obtained here for chains with those obtained previously for the three-center bond.

2.6.2 Rings



The Hamiltonian looks very similar to that of a linear chain. The only difference is that the first atom in the ring is connected to the last one. Thus there is an addition hopping parameter in the upper right and the lower left corner of the matrix.

$$\mathbf{H} = \begin{pmatrix} \epsilon_0 & t & 0 & 0 \cdots & t \\ t & \epsilon_0 & t & 0 \cdots & 0 \\ 0 & t & \epsilon_0 & t \cdots & 0 \\ 0 & 0 & t & \epsilon_0 \cdots & 0 \\ \vdots & \vdots & & & \\ t & 0 \cdots & 0 & t & \epsilon_0 \end{pmatrix}$$

Boundary conditions

In contrast to the linear chain, the ring has full translational symmetry, which is not even destroyed by the boundaries. The boundary conditions can again be expressed by the first and the last line of the Schrödinger equation. However, we can also treat the ring as an infinite chain with the requirement, that the wave function is periodic, that is $c_{N+1} = c_1$. We use the ansatz

$$c_j = \mathrm{e}^{ik\Delta j}$$

The boundary condition $c_{N+1} = c_1$ requires that

$$c_{N+1} = \mathrm{e}^{ik\Delta(N+1)} = c_1 = \mathrm{e}^{ik\Delta} \qquad \Rightarrow \qquad \mathrm{e}^{ik\Delta N} = 1 \qquad \Rightarrow \qquad k_n \Delta N = 2\pi n$$
$$\Rightarrow k_n = \frac{2\pi}{N\Delta} n \tag{2.33}$$

Exclude trivial quantum numbers

Two wave functions for two quantum numbers n and n' are identical, if

$$\begin{aligned} \mathrm{e}^{ik_n\Delta j} &= \mathrm{e}^{ik_{n'}\Delta j} \qquad \text{for all integer } j \\ \Rightarrow \mathrm{e}^{i(k_n - k_{n'})\Delta j} &= 1 \qquad \text{for all integer } j \\ \Rightarrow k_{n'} &= k_n + \frac{2\pi}{\Delta} q \qquad \text{with integer } q \\ \Rightarrow n' &= n + qN \end{aligned}$$

States with quantum numbers that differ by N are identical. Therefore we limit wave vectors k_n to the interval $]-\frac{\pi}{2}, \frac{\pi}{2}]$. Only one of the values at the boundaries may be included.

STATES OF A RING

Thus we obtain the wave functions and the energy eigenvalues as

$$\begin{aligned} |\psi_n\rangle &= \sum_j |\chi_j\rangle \mathrm{e}^{ik_n\Delta j} \frac{1}{\sqrt{N}} \\ \epsilon_n &= \epsilon_0 + 2t\cos(k_n\Delta) \\ k_n &= \frac{2\pi}{N\Delta} n \qquad \text{with} \qquad n = -\frac{N}{2} + 1, \ldots, \frac{N}{2} \end{aligned}$$

The resulting k-values lie in the interval $]\frac{\pi}{\Delta}, \frac{\pi}{\Delta}]$.

The spacing in k-space is approximately twice as large as in the linear chain with the same number of beads, but values with positive and negative k_n are allowed.

Each pair of degenerate wave functions can be transformed into a pair of real wave functions, of which one is a sinus and the other is a cosine.

Geometrical construction of the eigenstates

There is an interesting and easy way to memorize construction for the energy levels of a ring. The construction is demonstrated in Fig. 2.5. If one inscribes an equilateral polygon corresponding to the ring structure into a sphere, one can easily obtain the energy levels of the ring structure as vertical position of the corners.

The construction follows directly from the quantization condition

$$k_n = \frac{2\pi}{N\Delta} n \qquad \text{for} \qquad n = 0, 1, 2, \ldots, N$$

and the dispersion relation

$$\epsilon(k) = \epsilon_0 + 2t\cos(k\Delta)$$

Let us interpret $\phi := k\Delta$ as an angle. The allowed values of the angle are $\phi_n = \frac{2\pi}{N} n$ and the allowed energy values are

$$\epsilon_n = \epsilon_0 + 2t\cos(\phi_n)$$

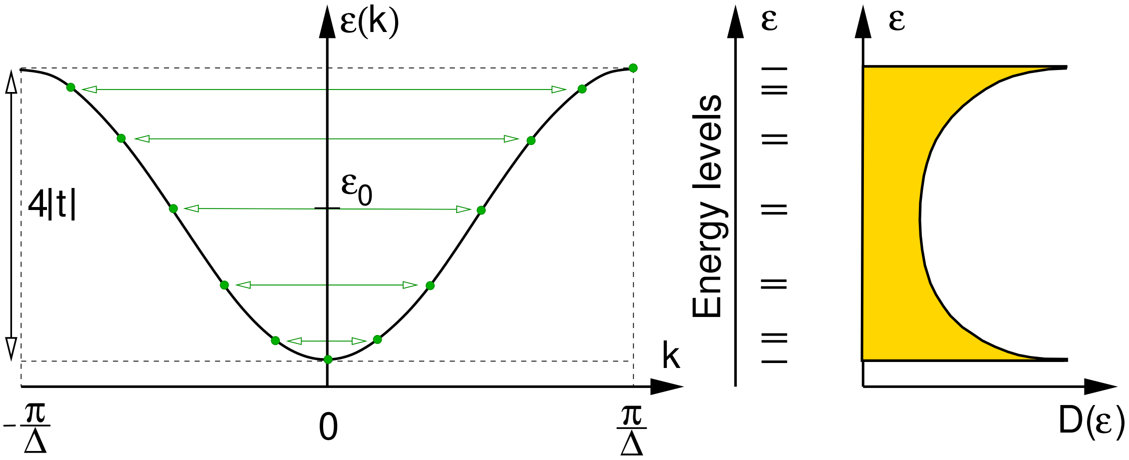


Fig. 2.4: Dispersion relation of the ring. The green points correspond to the energy levels of a ring with 14 atoms. In the middle figure we show the energy levels, and on the right the schematic density of states $D(\epsilon)$ for the infinite chain.

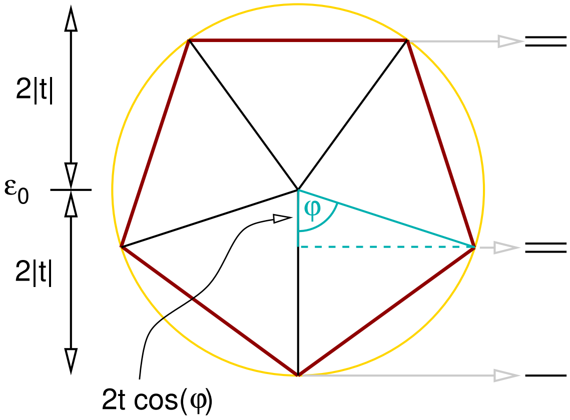


Fig. 2.5: Geometrical construction for the energy levels of a symmetric ring structure with n corners. The angles are $\phi = k_n\Delta = \frac{2\pi n}{N}$. Note that t is assumed to be negative.. With this assumption, the polygon stands on its top. The vertical position of the sphere center is at ϵ_0, and the radius is $2|t|$. The vertical positions of the corners are at the energy level positions $\epsilon_0 + 2t\cos(k_n\Delta)$

The cosine is the ratio of the adjacent leg[10] to the hypotenuse of a rectangular triangle[11]. Thus if the adjacent leg points from a point at height ϵ_0 straight down (because t is negative), the end point is the projection of the hypotenuse on the vertical axis. The hypotenuse shall have the length $2|t|$ and the end points lie on a circle with radius $2|t|$. Since the allowed values of the angles, are equispaced, and $\phi = 0$ is an allowed angle, the end-points of the hypotenuses form the corners of an equilateral polygon inscribed into the circle. The projection of the corners on the vertical axis gives us the allowed eigenvalues.

---

[10] adjacent leg is called “Ankathete” in german.

[11] A rectangular triangle is called “rechtwinkliges Dreieck” in german

Chapter 3

The use of symmetry

3.1 Introduction

Exploiting the symmetry allows one to break down a large eigenvalue problem into many smaller ones. The complexity of a matrix diagonlization grows with the third power of the matrix dimension. Therefore, it is much easier to diagonalize many smaller matrices than one big one, that is composed of the smaller ones. If one manages to break down the problem into ones with only one, two or three orbitals, we may use the techniques from the previous chapter to diagonalize the Hamiltonian. This often allows one to obtain an educated guess of the wave functions without much computation.

As shown below, the eigenstates of the Hamiltonian are also eigenstates of the symmetry operator, if the Hamiltonian has a certain symmetry. (For degenerate states this is not automatically so, but a transformation can bring them into the desired form.) Furthermore, we will show the following statement, which will be central to this section:

BLOCK DIAGONALIZATION BY SYMMETRY

The Hamilton matrix elements between eigenstates of a symmetry operator with different eigenvalues vanish.

In other words:

In a basis of symmetry eigenstates, the Hamiltonian is **block diagonal**. All states in a given block agree in all symmetry eigenvalues.

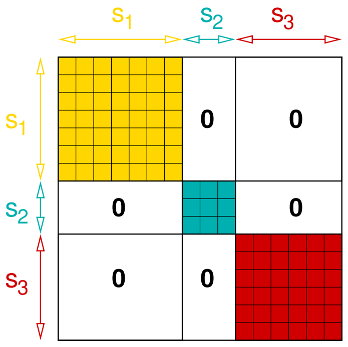


Fig. 3.1: Block form of a Hamiltonian by using symmetry eigenstates with symmetry eigenvalues s_1, s_2, s_3.

3.2 Symmetry and quantum mechanics

Here, I will revisit the main **symmetry** arguments discussed in chapter 9 of ΦSX:Quantum Physics. This is a series of arguments that is good to keep in mind.

1. Definition of a **transformation operator**. An operator $\hat{S}$ can be called a transformation, if it conserves the norm for every state.
$$\forall_{|\psi\rangle} \qquad \langle\psi|\psi\rangle \overset{|\phi\rangle=\hat{S}|\psi\rangle}{=} \langle\phi|\phi\rangle \tag{3.1}$$
2. Every transformation operator is unitary, that is $\hat{S}^\dagger\hat{S}=1$
$$\forall_{|\psi\rangle} \quad \langle\psi|\hat{S}^\dagger\hat{S}|\psi\rangle = \langle\psi|\psi\rangle \qquad \Rightarrow \qquad \hat{S}^\dagger\hat{S}=1 \tag{3.2}$$
3. Definition of a **symmetry**: A system is symmetric under the transformation $\hat{S}$, if, for any solution $|\Psi\rangle$ of the Schrödinger equation describing that system, also $\hat{S}|\Psi\rangle$ is a solution of the same Schrödinger equation. That is, if
$$\Big(\quad i\hbar\partial_t|\Psi\rangle = \hat{H}|\Psi\rangle \quad\Big) \overset{|\Phi\rangle:=\hat{S}|\Psi\rangle}{\Rightarrow} \Big(\quad i\hbar\partial_t|\Phi\rangle = \hat{H}|\Phi\rangle \quad\Big) \tag{3.3}$$
4. A unitary operator $\hat{S}$ is a symmetry operator of the system, if
$$[\hat{H},\hat{S}]_- = 0 \tag{3.4}$$
Proof: Let $|\Psi\rangle$ be a solution of the Schrödinger equation and let $|\Phi\rangle = \hat{S}|\Psi\rangle$ be the result of a symmetry transformation.
$$\begin{aligned} i\hbar\partial_t|\Phi\rangle &= \hat{H}|\Phi\rangle \\ \overset{|\Phi:=\hat{S}|\Psi\rangle}{\Rightarrow} \; i\hbar\partial_t\hat{S}|\Psi\rangle &= \hat{H}\hat{S}|\Psi\rangle \\ \overset{\partial_t\hat{S}=0}{\Rightarrow} \; i\hbar\partial_t|\Psi\rangle &= \hat{S}^{-1}\hat{H}\hat{S}|\Psi\rangle \\ \overset{i\hbar\partial_t|\Psi\rangle=\hat{H}|\Psi\rangle}{\Rightarrow} \; \hat{H}|\Psi\rangle &= \hat{S}^{-1}\hat{H}\hat{S}|\Psi\rangle \\ \underbrace{\left(\hat{H}\hat{S}-\hat{S}\hat{H}\right)}_{[\hat{H},\hat{S}]_-}|\Psi\rangle &= 0 \end{aligned}$$
Because this equation holds for any solution of the Schrödinger equation, it holds for any wave function, because any function can be written as superposition of solutions of the Schrödinger equation. (The latter form a complete set of functions.) Therefore
$$[\hat{H},\hat{S}]_- = 0$$
Thus, one usually identifies a symmetry by working out the commutator with the Hamiltonian.
5. The matrix elements of the Hamilton operator between two eigenstates of the symmetry operator with different eigenvalues vanish. That is
$$\Big(\quad \hat{S}|\Psi_s\rangle = |\Psi_s\rangle s \quad\wedge\quad \hat{S}|\Psi_{s'}\rangle = |\Psi_{s'}\rangle s' \quad\wedge\quad s\neq s' \quad\Big) \quad\Rightarrow\quad \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle = 0 \tag{3.5}$$
Proof: In the following we will need an expression for $\langle\psi_s|\hat{S}$, which we will work out first:
    - We start showing that the absolute value of an eigenvalue of a unitary operator is equal to one, that is $s=\mathrm{e}^{i\phi}$ where ϕ is real. With an eigenstate $|\psi_s\rangle$ of $\hat{S}$ with eigenvalue s, we obtain
    $$s^*\langle\psi_s|\psi_s\rangle s \overset{\hat{S}|\psi_s\rangle=|\psi_s\rangle s}{=} \langle\hat{S}\psi_s|\hat{S}\psi_s\rangle = \langle\psi_s|\underbrace{\hat{S}^\dagger\hat{S}}_{\hat{1}}|\psi_s\rangle = \langle\psi_s|\psi_s\rangle$$
    $$\Rightarrow \qquad |s|=1 \tag{3.6}$$

- Next we show that the eigenvalues of the hermitian conjugate operator $\hat{S}^\dagger$ of a unitary operator $\hat{S}$ are the complex conjugates of the eigenvalues of $\hat{S}$.

$$\hat{S}^\dagger|\psi\rangle \overset{\hat{S}^\dagger\hat{S}=1}{=} \hat{S}^{-1}|\psi_s\rangle \overset{\hat{S}|\psi_s\rangle=|\psi_s\rangle s}{=} |\psi_s\rangle s^{-1} = |\psi_s\rangle\frac{s^*}{ss^*} \overset{|s|=1}{=} |\psi_s\rangle s^* \tag{3.7}$$

- Now, we are ready to show that the matrix elements of the Hamilton operator between two eigenstates of the symmetry operator with different eigenvalues vanish.

  We will use

$$\langle\psi_s|\hat{S} = s\langle\psi_s| \tag{3.8}$$

which directly follows from Eq. 3.7.

$$\begin{aligned}
0 \quad & \overset{[\hat{H},\hat{S}]_-=0}{=} \quad \langle\Psi_s|[\hat{H},\hat{S}]_-|\Psi_{s'}\rangle = \langle\Psi_s|\hat{H}\hat{S}|\Psi_{s'}\rangle - \langle\Psi_s|\hat{S}\hat{H}|\Psi_{s'}\rangle \\
& \overset{\hat{S}|\Psi_s\rangle=|\Psi_s\rangle s,\text{Eq. }3.8}{=} \quad \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle s' - s\langle\Psi_s|\hat{H}|\Psi_{s'}\rangle = \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle(s'-s) \\
& \overset{s\neq s'}{\Rightarrow} \quad \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle = 0
\end{aligned} \tag{3.9}$$

6. .

CONSTRUCTION OF SYMMETRY EIGENSTATES

Eigenstates of a symmetry operator with a finite symmetry group, that is $\hat{S}^N = \hat{1}$ for a given eigenvalue s_α can be constructed from an arbitrary state $|\chi\rangle$ by superposition.

$$|\Psi_\alpha\rangle = \sum_{n=0}^{N-1} \hat{S}^n|\chi\rangle s_\alpha^{-n} \tag{3.10}$$

The eigenvalues s_α of the symmetry operator $\hat{S}$ are

$$s_\alpha = \mathrm{e}^{i\frac{2\pi}{N}\alpha} , \tag{3.11}$$

which follows from $|s_\alpha| = 1$ (Eq. 3.6) and $s_\alpha^N = 1$. The operation in Eq. 3.10 acts like a filter, that projects out all contributions from eigenvalues other than the chosen one. Therefore, the result may vanish.

Proof:

$$\begin{aligned}
\hat{S}|\Psi_\alpha\rangle &= \hat{S}\sum_{n=0}^{N-1}\hat{S}^n|\chi\rangle s_\alpha^{-n} = \left[\sum_{n=0}^{N-1}\hat{S}^{n+1}|\chi\rangle s_\alpha^{-(n+1)}\right] s_\alpha \\
&\overset{\hat{S}^N=\hat{1};s_\alpha^N=1}{=} \left[\sum_{n=0}^{N-1}\hat{S}^n|\chi\rangle s_\alpha^{-n}\right] s_\alpha = |\Psi_\alpha\rangle s_\alpha
\end{aligned}$$

With what has been discussed above, we have shown that the Hamilton operator is block diagonal in a representation of eigenstates of its symmetry operators. The eigenstates of the Hamilton operator can be obtained for each block individually. For us, it is more important that a wave function that starts out as an eigenstate of a symmetry operator to a given eigenvalue, will always remain an eigenstate to the same eigenvalue. In other words, the eigenvalue of the symmetry operator is a conserved quantity. (Note, that symmetry operators usually have complex eigenvalues)

The eigenvalues of the symmetry operators are the **quantum numbers**.

3.3 Symmetry eigenstates of the hydrogen molecule

Let us consider the hydrogen molecule and let us only consider the two s-orbitals. The hydrogen molecule is symmetric with respect to a mirror plane in the bond center.

We orient the hydrogen molecule in z-direction and place the bond-center into the origin. The mirror operation $\hat{S}$ about the bond center has the form

$$\hat{S}\psi(x, y, z) = \psi(x, y, -z)$$

or in bra-ket notation

$$\langle x, y, z|\hat{S}|\psi\rangle = \langle x, y, -z|\psi\rangle$$

If we apply the mirror operation twice, we obtain the original result. Therefore, $\hat{S}^2 = \hat{1}$ is the identity. For an eigenstate

$$\hat{S}|\psi\rangle = |\psi\rangle s$$

we obtain

$$\hat{S}^2|\psi\rangle = |\psi\rangle s^2 \overset{\hat{S}^2=\hat{1}}{=} |\psi\rangle 1$$

Hence we obtain $s^2 = 1$, which has two solutions, namely $s = 1$ and $s = -1$.

We can now construct symmetrized orbitals out of our basis functions:

- for the eigenvalue $s = +1$ we obtain the symmetric state, namely
$$|\chi'_1\rangle = \frac{1}{\sqrt{2}}\left(|\chi_1\rangle + |\chi_2\rangle\right)$$
- for the eigenvalue $s = -1$ we obtain the anti-symmetric state, namely
$$|\chi'_2\rangle = \frac{1}{\sqrt{2}}\left(|\chi_1\rangle - |\chi_2\rangle\right)$$

The factor $\frac{1}{\sqrt{2}}$ has been added to ensure that also the new orbitals are orthonormal.

In matrix-vector notation we may write[1]

$$\begin{pmatrix} |\chi'_1\rangle \\ |\chi'_2\rangle \end{pmatrix} = \begin{pmatrix} |\chi_1\rangle \\ |\chi_2\rangle \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Now we can transform the Hamiltonian into the new basis set. In terms of the atomic orbitals the Hamilton matrix is

$$\mathbf{H} = \begin{pmatrix} \bar{\epsilon} & t \\ t & \bar{\epsilon} \end{pmatrix}$$

The Hamiltonian in the basis of the new orbitals is

$$\mathbf{H}' = \begin{pmatrix} \langle\chi'_1|\hat{H}|\chi'_1\rangle & \langle\chi'_1|\hat{H}|\chi'_2\rangle \\ \langle\chi'_2|\hat{H}|\chi'_1\rangle & \langle\chi'_2|\hat{H}|\chi'_2\rangle \end{pmatrix} = \mathbf{U}^\dagger\mathbf{H}\mathbf{U} = \begin{pmatrix} \bar{\epsilon} + t & 0 \\ 0 & \bar{\epsilon} - t \end{pmatrix}$$

In this case we have already diagonalized the matrix. By introducing a basisset made of symmetrized orbitals the Hamiltonian became block diagonal with two 1×1 blocks.

---

[1] In the following expression one usually write the matrix to the left and the vector to the right. I use the convention that, in an expression that is a ket-vector, the ket stands always on the left side and the matrix, on the right side. From one notation to the other, it is necessary to transpose the matrix accordingly. For bra-expression the bra-vector stands on the right side, and the matrix on the left. My notation has the advantage that the order of the individual terms remains unchanged, when a bra and a ket is combined to a matrix element (bra-ket).

3.4 Symmetry eigenstates of a main group dimer

Let us now consider the dimeric main group dimers. In addition to the two s-orbitals used for hydrogen, we need to include also the p orbitals in the basis. The main group dimer is, in a way, the prototype for a bond between two main group atoms.

As for the hydrogen molecule we will follow a sequence of steps:

1. choose a basisset
2. determine the symmetry of the molecule
3. select a subset of symmetry operations. Preferably, one selects operations that commutate with each other. In that case they can be treated as independent.
4. For each set of eigenvalues of the symmetry operators, form the corresponding symmtrized orbitals from the basis orbitals.

This procedure produces groups of orbitals, that block diagonalize the Hamiltonien. This is all one can do with symmetry operations. What remains to be done is to diagonalize the remaining sub-blocks of the Hamiltonian. This can be done in several ways, short of doing an ab-initio calculation.

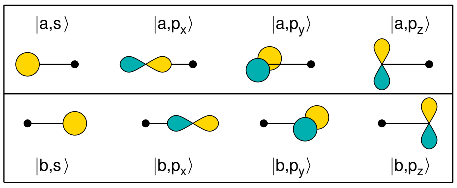
- One can set up a parameterized Hamiltonian for the subblocks and diagonalize them. If the sub groups only contain less than three orbitals, they can be diagonalized directly. If more orbitals are involved, one better uses the computer.
- One can do the first steps of an iterative diagonalization by hand. This gives only approximate answers, but often provides most insight.

Choose a basis

The first step is to choose a set of basis states. We start out with atom-centered orbitals on the two sites, denoted by *a* (the left one) and *b* (the right one).

The angular dependence is characterized by a "real" spherical harmonic ($s, p_x, p_y, p_z, d_{x^2-y^2}, \ldots$). The first few have the form

$$Y_s(\vec{r}) = \frac{1}{\sqrt{4\pi}} \quad ; \quad Y_{p_x}(\vec{r}) = \sqrt{\frac{3}{4\pi}}\frac{x}{|\vec{r}|} \quad ; \quad Y_{p_y}(\vec{r}) = \sqrt{\frac{3}{4\pi}}\frac{y}{|\vec{r}|} \quad ; \quad Y_{p_z}(\vec{r}) = \sqrt{\frac{3}{4\pi}}\frac{z}{|\vec{r}|}$$



The angular dependence is illustrated as lobes wich are a schematic drawing of the polar-coordinate representation of the real spherical harmonics.

The fat dot represents an atom for which none of its atomic orbitals contribute to the wave function.

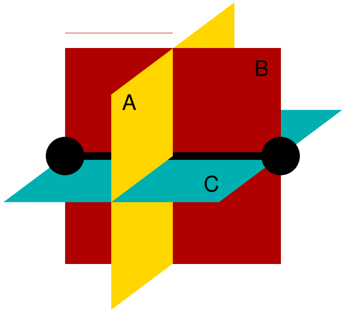
Symmetry operations of the molecule

We again by determining the symmetry of the molecule.

- mirror plane perpendicular to the bond through the bond center

- mirror planes with the bond in the mirror plane
- continuous rotational symmetry about the bond
- two fold rotational symmetry about any axis perpendicular to the bond, with the axis passing through the bond center
- point inversion about the bond center

We only need to pick out a few.[2]
Since we are already familiar with mirror operations, let us pick the three mirror planes.

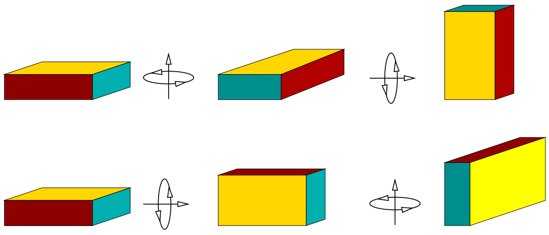


One should select a subset of symmetry operations that commutate with each other. Symmetry operations commutate, if the order, in which they are performed, does not affect result.

Example for non-commutating operations

An example for two operations that do not commutate is given in Fig. 3.2 on p. 38.

$$\underbrace{\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{R_z(90^\circ)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}_{R_x(90^\circ)} = \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{R_{(111)}(90^\circ)}$$



$$\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}^{R_x(90^\circ)} \overbrace{\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{R_z(90^\circ)} = \overbrace{\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}}^{R_{(1\bar{1}1)}(90^\circ)}$$

Fig. 3.2: Demonstration of two non-commutative rotations. The result (top) of a 90° rotation about the z-axis (vertical) followed by a 90° about the x-axis (horizontal, in the paper plane) differs from the result (bottom) obtained when the same operations are performed in reversed order. $R_{(111)}(90^\circ)$ denotes a 90° rotation about the diagonal pointing along $(1,1,1)$. $R_{(1\bar{1}1)}(90^\circ)$ denotes a 90° rotation about the $(1,-1,1)$ axis

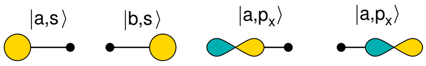
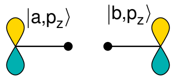
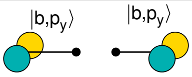
---

[2] The result does not become wrong, if we do not exploit all symmetries. If we miss out some, the blocks of the resulting Hamiltonian may be larger, which makes the diagonalization more cumbersome.

Eigenstates of the symmetry operators

Our basis orbitals are an s-orbital on each atom and three p-orbitals on each atom. We group them according to the eigenvalues for the three mirror planes A, B, C.

We begin by grouping them according to the eigenvalues of B and C, because the orbitals we have chosen are already eigenstates of these operations. The $B+$ indicates the states that are symmetric with respect to the mirror plane B and $B-$ indicates the states that are antisymmetric with respect to the mirror plane B. We use the analogous notation for the mirror planes A and C. On the right hand side, we list those orbitals that have the corresponding behavior with respect to the two mirror planes.

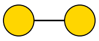
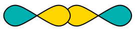
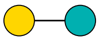
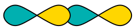
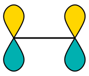
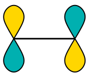
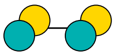
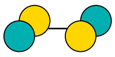
B+	C+	 |a,s⟩ |b,s⟩ |a,p_x⟩ |a,p_x⟩
B+	C−	
B−	C+	
B−	C−	

Now we form the symmetrized states with respect to A using Eq. 3.10 from p. 35. Note that we need not consider any mixing between states from sets with different symmetry eigenvalues.

$$|\Psi_\alpha\rangle \overset{\text{Eq. 3.10}}{=} \sum_{n=1}^{N-1} \hat{S}^n|\chi\rangle s_\alpha^{-n} \tag{3.12}$$

We start with $|a,s\rangle$ and form $\hat{S}_A|a,s\rangle + |b,s\rangle$. Thus the symmetric state with respect to the mirror plane A is $\frac{1}{\sqrt{2}}\big(|a,s\rangle = |b,s\rangle\big)$ and the antisymmetric state is $\frac{1}{\sqrt{2}}\big(|a,s\rangle - |b,s\rangle\big)$. Starting from $|b,s\rangle$ gives the same states. The state $|a,p_x\rangle$ transforms into $\hat{S}_A|a,p_x\rangle = -|b,p_x\rangle$. Therefore the symmetric state is $\frac{1}{\sqrt{2}}\big(|a,p_x\rangle - |b,p_x\rangle\big)$ and the antisymmetric state is $\frac{1}{\sqrt{2}}\big(|a,p_x\rangle + |b,p_x\rangle\big)$. Note, that a state with a positive sign is not automatically a symmetric state. No new states are obtained starting from $|b,p_x\rangle$.

An important cross check is to verify that the number of orbitals in each group is the same before and after symetrization.

A+	B+	C+	$\frac{1}{\sqrt{2}}\big(\lvert a,s\rangle+\lvert b,s\rangle\big)$   $\frac{1}{\sqrt{2}}\big(\lvert a,p_x\rangle-\lvert b,p_x\rangle\big)$
A−	B+	C+	$\frac{1}{\sqrt{2}}\big(\lvert a,s\rangle-\lvert b,s\rangle\big)$   $\frac{1}{\sqrt{2}}\big(\lvert a,p_x\rangle+\lvert b,p_x\rangle\big)$
A+	B+	C−	 $\frac{1}{\sqrt{2}}\big(\lvert a,p_z\rangle+\lvert b,p_z\rangle\big)$
A−	B+	C−	 $\frac{1}{\sqrt{2}}\big(\lvert a,p_z\rangle-\lvert b,p_z\rangle\big)$
A+	B−	C+	 $\frac{1}{\sqrt{2}}\big(\lvert a,p_y\rangle+\lvert b,p_y\rangle\big)$
A−	B−	C+	 $\frac{1}{\sqrt{2}}\big(\lvert a,p_y\rangle-\lvert b,p_y\rangle\big)$
A+	B−	C−	
A−	B−	C−	

Instead of solving a 8-dimensional eigenvalue problem, we only need to determine two 2-dimensional eigenvalue problems, a tremendous simplification.

The two-dimensional problem can be estimated graphically. They correspond to the non-degenerate problem of the two-center bond, discussed in section 2.3.

3.4.1 Approximate diagonalization

Up to now we have done all that can be done for the main-group dimer using symmetry alone. Now we need to diagonalize the remaining sub-blocks of the symmetry eigenstates. Here, I will show how one can get an idea of the eigenstates graphically, that is without any calculation. The results will be approximate. Later we will see how to set up a parameterized Hamiltonian, and do the calculation for a simple **tight-binding model**.

There are two problems to be solved:

- diagonalize the remaining sub-blocks containing more than one orbital
- determine the relative position of the orbitals

3.4.2 σ bonds and antibonds

The terms σ and π bond will be explained below. The σ bonds and antibonds result from the 2×2 blocks with orbitals that have cylinder symmetry.

Let us consider once the s orbitals on both atoms. As known from the section on the two-center boind, they form a bonding and an antibonding orbital. The bonding orbital is symmetric with respect to mirror operation at the bond-center plane, and the antibonding orbital is antisymmetric. This is

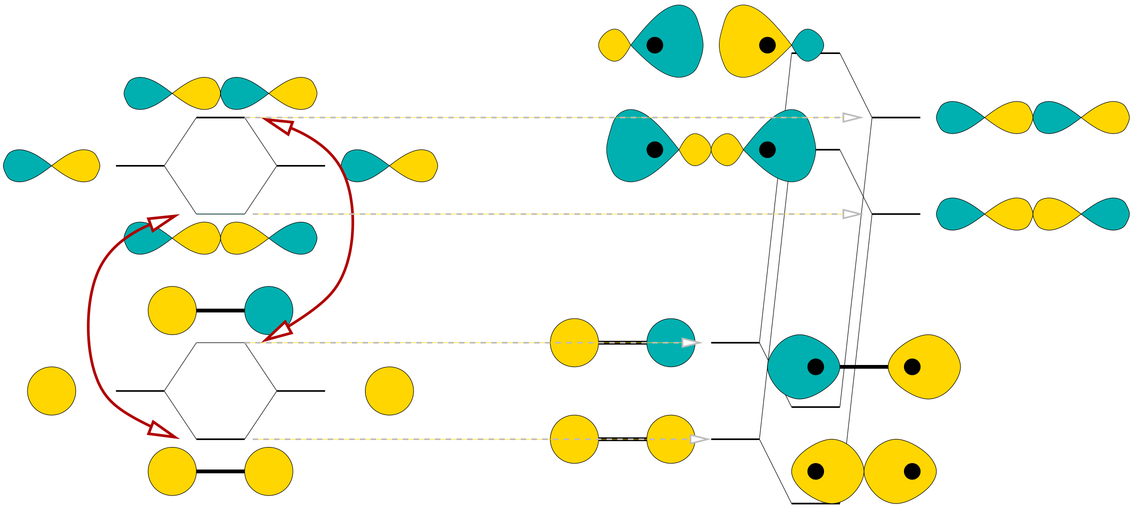


Fig. 3.3: Schematic drawing of the hybridization of the σ states of a dimeric molecule. The vertical axis corresponds to the energy of the orbitals

the information we know also from the symmetry consideration, but now we also know that the energy levels are centered at the "atomic s orbital energy".

Similarly we form the bonding and antibonding p-orbitals, that again form a symmetric (bonding) and an antisymmetric (antibonding) state.

From our symmetry considerations we know that only the bonding orbitals interact with each other and the two antibonding orbitals interact with each other. This is indicated by the red arrows in figure 3.3. Because they belong to different symmetry eigenvalues, the bonding orbitals do not interact with the antibonding orbitals.

The hybridization of the two bonding orbitals is analogous to the two-center bond in the non-degenerate case. Note that the model of a two center bond gave us a general recipy on diagonalizing 2×2 Hamiltonians.

The two states of that model need not be orbitals on the two bonding partners as in section 2.3, but they can be any pair of orbitals that have a Hamilton matrix element between them.

The lower bonding level will have mostly s-character and it will lie even below the energy the pure s-type orbital. The bonding p-orbital mixes into the s-type bond orbital so as to enhance the orbital weight in the center of the bond, and to attenuate the wave function in the back bond. (pointing away from the neighbor). This mixing of s and p-orbital forms so-called hybrid orbitals that willl be discussed later.

The higher bonding orbital will lie above the p-type bonding orbital, and it will have predominantly p-type character. However the s-type bonding orbital mixes in, but now in the opposite compared to the bond orbital discussed above. Now the orbital in the middle of the bond is attenuated and enhanced in the back bond. This weakening of the bond explains that the p-type bond orbital shifts up in energy.

Similarly we can analyze the interaction between the two antibonding orbitals. Only here the s-type antibond is weakened, so that it shifts towards lower energy and the p-type antibond is made stronger, shifting it further up in energy.

3.4.3 π bonds and antibonds

The 1×1 blocks form the π bonds and antibonds shown in Fig. 3.4. Using what we learned from the symmetric two-center bond, we know that bonding and antibonding orbitals are centered around the “atomic p-level”. From the rotational symmetry of the bond, we can infer that the two bonding orbitals are degenerate and that the two antibonding orbitals are degenerate, too.

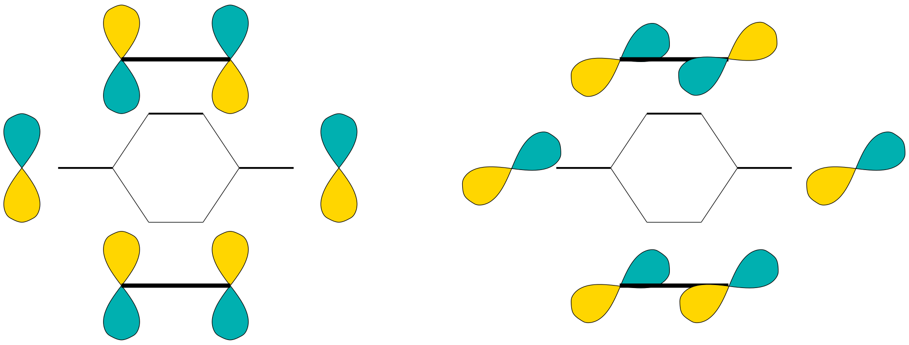


Fig. 3.4: Schematic drawing of the hybridization of the π-states of a dimeric molecule. The vertical axis corresponds to the energy of the orbitals

Finally we can place the orbitals into one diagram shown in Fig.3.5. The center of the π orbitals will lie below the two upper σ bonds, because the latter have been shifted up by the hybridization with the s orbitals.

The splitting between the π orbitals will be smaller than that of the σ orbitals, because the latter point towards each other and therefore have a larger overlap and Hamilton matrix element.

The order of the orbitals is not completely unique and,in particular, the σ-p-bond and π-bond can interchange as one goes through a period in the periodic table.

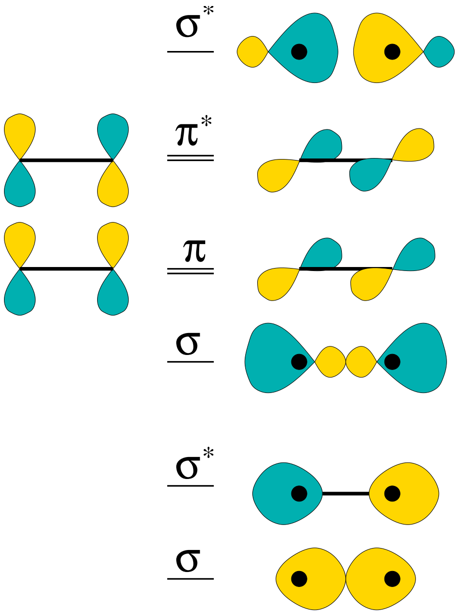


Fig. 3.5: Schematic energy level diagram of a main group dimer

3.5 Parameterized Hamiltonian of the main group dimer

In the following, I will demonstrate the the steps discussed so far for a parameterized Hamiltonian. In its symmetrized form, such a parameterized Hamiltonian allows us to diagonalize the remaining blocks of the Hamiltonian.

3.5.1 Notion of σ, π, δ-states

The states of a dimer are classified as to σ and π states. This notation is analogous to the classification of atomic orbitals according to their angular momentum as s, p, d, f-states. For systems with a single rotation axis, we classify the states according to their angular momentum about the bond axis as σ, π, δ-states.

- A state that is rotationally symmetric is called a σ-state. A sigma state has the angular momentum quantum number is $m = 0$ for rotation about the bond axis.

- A state that changes its sign twice when moving in a circle about the bond axis is a π state. A π-state has the angular momentum quantum number is $m = 1$ for rotation about the bond axis.

- A state that changes its sign four times when moving in a circle about the bond axis, is a δ-state. A δ-state has the angular momentum quantum number is $m = 2$ for rotation about the bond axis.

The notation carries a meaning beyond that of symmetry. σ bonds and anti-bonds are typically the strongest. The π-bonds are weaker and the δ-bonds are nearly negligible.

Antibonding orbitals are distinguished from the bond-orbitals by a star such as σ^*, pronounced **"sigma-star"**. The best is to inspect figure 3.5.

3.5.2 Slater-Koster tables

Slater and Koster[3] have formed the basis of today's **empirical tight-binding methods**[3]. They showed that the Hamiltonian and overlap matrix can be built up easily using their **Slater-Koster tables**[3], if one knows the matrix elements of Hamilton and overlap matrix for certain symmetric combinations along a bond. These symmetric combinations are sketched in Fig. 3.6.

If the matrix elements are furthermore known as function of distance one can also investigate the changes of the electronic structure, respectively estimate the forces.

3.5.3 Hopping matrix elements of Harrison

Harrison (*"The physics of solid state chemistry"*, Walter A. Harrison, Advances in Solid state Physics 17, p135)[8] has shown that the matrix elements can be reasonably fit by simple expressions.

- Harrison chose the orbital energies equal to the atomic energy levels obtained with the Hartree Fock approximation.

- The parameters $t_{ss\sigma}, t_{sp\sigma}, t_{pp\sigma}, t_{pp\pi}$ for the hopping parameters can be derived by fitting the

---

[3] The kind of empirical models, where only nearest neighbor matrix elements of Hamiltonian are considered, and overlap is set to unity is called tight binding models. However the definition is not clear cut and there are many variants and extensions that run with the same name.

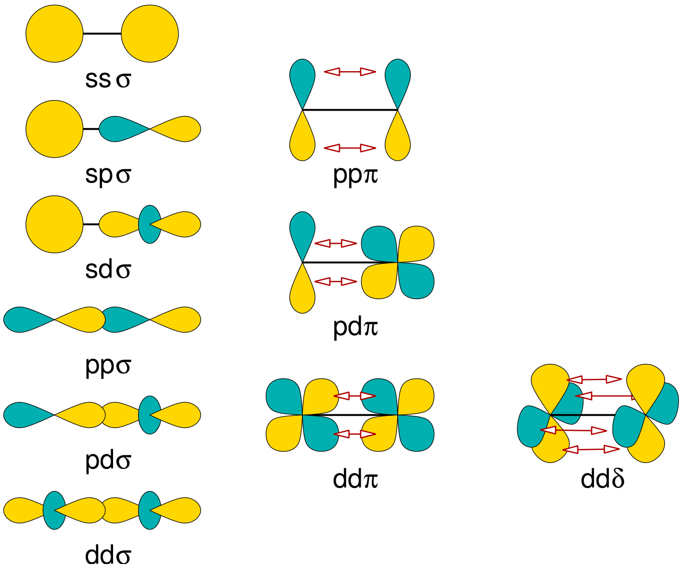


Fig. 3.6: Non-vanishing tight-binding matrix elements along a bond and their naming.

free electron bands with tight-binding matrix elements,[9]. The values are

$$
\begin{aligned}
t_{ss\sigma} &= 1.40\frac{\hbar^2}{m_e d^2} \\
t_{sp\sigma} &= -1.84\frac{\hbar^2}{m_e d^2} \\
t_{pp\sigma} &= -3.24\frac{\hbar^2}{m_e d^2} \\
t_{pp\pi} &= 0.81\frac{\hbar^2}{m_e d^2}
\end{aligned}
\tag{3.13}
$$

where d is the interatomic distance.

Further reading can be found in [10, 11].

3.5.4 Parameterized Hamiltonian of the main group dimer

I will demonstrate how to parameterize a Hamiltonian and how to work out the energy levels in terms of these parameters. I will use the example of a dimeric molecule as an example.

Let us first consider the energies of the orbitals

$$
\begin{aligned}
\bar{\epsilon}_s &= \langle a, s|\hat{H}|a, s\rangle = \langle b, s|\hat{H}|b, s\rangle \\
\bar{\epsilon}_p &= \langle a, p_x|\hat{H}|a, p_x\rangle = \langle a, p_y|\hat{H}|a, p_y\rangle = \langle a, p_z|\hat{H}|a, p_z\rangle \\
&= \langle b, p_x|\hat{H}|b, p_x\rangle = \langle b, p_y|\hat{H}|b, p_y\rangle = \langle b, p_z|\hat{H}|b, p_z\rangle
\end{aligned}
$$

Usually one thinks here of the atomic orbitals.

For the off-site terms, we will use the, admittedly rather crude, approximations of choosing the Hamilton matrix elements proportional to the overlap matrix elements and then to ignore the deviations from orthonormality. Due to the rotational symmetry about the molecular axis, only certain matrix elements are nonzero.[4] This rotational symmetry is another simplifying assumption.

---

[4] Symmetry blockdiagonalizes the Hamiltonian. Therefore many matrix elements of the Hamiltonian become zero in a basis of symmetry eigenstates.

$$\begin{aligned}
t_{ss\sigma} &= \langle a,s|\hat{H}|b,s\rangle \\
t_{sp\sigma} &= \langle a,s|\hat{H}|b,p_x\rangle \\
t_{pp\sigma} &= \langle a,p_x|\hat{H}|b,p_x\rangle \\
t_{pp\pi} &= \langle a,p_y|\hat{H}|b,p_y\rangle = \langle a,p_z|\hat{H}|b,p_z\rangle
\end{aligned}$$

If we describe the states as a vector with basisstates in the order ($|a,s\rangle$, $|a,p_x\rangle$, $|a,p_y\rangle$, $|a,p_z\rangle$, $|b,s\rangle$, $|b,p_x\rangle$, $|b,p_y\rangle$, $|b,p_z\rangle$), the Hamilton matrix has the form

$$\boldsymbol{H} \hat{=} \begin{pmatrix}
\epsilon_s & 0 & 0 & 0 & t_{ss\sigma} & t_{sp\sigma} & 0 & 0 \\
0 & \epsilon_p & 0 & 0 & -t_{sp\sigma} & t_{pp\sigma} & 0 & 0 \\
0 & 0 & \epsilon_p & 0 & 0 & 0 & t_{pp\pi} & 0 \\
0 & 0 & 0 & \epsilon_p & 0 & 0 & 0 & t_{pp\pi} \\
t_{ss\sigma} & -t_{sp\sigma} & 0 & 0 & \epsilon_s & 0 & 0 & 0 \\
t_{sp\sigma} & t_{pp\sigma} & 0 & 0 & 0 & \epsilon_p & 0 & 0 \\
0 & 0 & t_{pp\pi} & 0 & 0 & 0 & \epsilon_p & 0 \\
0 & 0 & 0 & t_{pp\pi} & 0 & 0 & 0 & \epsilon_p
\end{pmatrix} \tag{3.14}$$

Symmetrized orbitals

We work out the Hamiltonian in the basis of eigenstates of the three symmetry planes, that we have determined in the beginning of section 3.4. These are

$$\begin{pmatrix} |\phi_1\rangle \\ |\phi_2\rangle \\ |\phi_3\rangle \\ |\phi_4\rangle \\ |\phi_5\rangle \\ |\phi_6\rangle \\ |\phi_7\rangle \\ |\phi_8\rangle \end{pmatrix} = \begin{pmatrix}
\frac{1}{\sqrt{2}}(|a,s\rangle + |b,s\rangle) \\
\frac{1}{\sqrt{2}}(|a,p_x\rangle - |b,p_x\rangle) \\
\frac{1}{\sqrt{2}}(|a,s\rangle - |b,s\rangle) \\
\frac{1}{\sqrt{2}}(|a,p_x\rangle + |b,p_x\rangle) \\
\frac{1}{\sqrt{2}}(|a,p_z\rangle + |b,p_z\rangle) \\
\frac{1}{\sqrt{2}}(|a,p_z\rangle - |b,p_z\rangle) \\
\frac{1}{\sqrt{2}}(|a,p_y\rangle + |b,p_y\rangle) \\
\frac{1}{\sqrt{2}}(|a,p_y\rangle - |b,p_y\rangle)
\end{pmatrix} \tag{3.15}$$

This can be done in two ways, which I demonstrate for one example, namely the matrix element between $\langle\phi_1| := \frac{1}{\sqrt{2}}(\langle a,s| + \langle b,s|)$ and $|\phi_2\rangle := \frac{1}{\sqrt{2}}(|a,p_x\rangle - |b,p_x\rangle)$

- We can represent the bra by a vector $\vec{c}_1 = (1,0,0,0,1,0,0,0)\frac{1}{\sqrt{2}}$ and the ket by the vector $\vec{c}_2 = (0,1,0,0,0,-1,0,0)\frac{1}{\sqrt{2}}$, so that

$$\begin{aligned}
\langle\phi_1| &= \sum_{i=1}^{8} \langle\phi_1|\chi_i\rangle\langle\chi_i| = \sum_{i=1}^{8} c^*_{i,1}\langle\chi_i| \\
|\phi_2\rangle &= \sum_{i=1}^{8} |\chi_i\rangle\langle\chi_i|\phi_2\rangle = \sum_{i=1}^{8} |\chi_i\rangle c_{i,2}
\end{aligned}$$

  Now, we multiply the Hamilton matrix from Eq. 3.14, which has elements $\langle\chi_i|\hat{H}|\chi_j\rangle$, with $\vec{c}^*_1$ from the left and with $\vec{c}_2$ from the right, which yields

$$\langle\phi_1|\hat{H}|\phi_2\rangle = \frac{1}{2}\left(-t_{sp\sigma} - t_{sp\sigma}\right) = -t_{sp\sigma}$$

- As an alternative we can also directly decompose the matrix elements of the symmetrized states into those of the original (not symmetrized) basis orbitals.

$$
\begin{aligned}
&\left[\frac{1}{\sqrt{2}}\left(\langle a,s| + \langle b,s|\right)\right]\hat{H}\left[\frac{1}{\sqrt{2}}\left(|a,p_x\rangle - |b,p_x\rangle\right)\right]\\
&=\frac{1}{2}\left(\underbrace{\langle a,s|\hat{H}|a,p_x\rangle}_{0}-\underbrace{\langle a,s|\hat{H}|b,p_x\rangle}_{t_{sp\sigma}}+\underbrace{\langle b,s|\hat{H}|a,p_x\rangle}_{-t_{sp\sigma}}-\underbrace{\langle b,s|\hat{H}|b,p_x\rangle}_{0}\right)\\
&=-t_{sp\sigma}
\end{aligned}
$$

In the basis of symmetry eigenstates from Eq. 3.15, the Hamiltonian obtains its block-diagonal form with two 2 × 2 blocks and and six eigenstates. The elements of this Hamilton matrix are $\langle\phi_i|\hat{H}|\phi_j\rangle$.

$$
\hat{H}\hat{=}\left(\begin{array}{cc|cc|c|c|c|c}
\epsilon_s+t_{ss\sigma} & -t_{sp\sigma} & 0 & 0 & 0 & 0 & 0 & 0\\
-t_{sp\sigma} & \epsilon_p+t_{pp\sigma} & 0 & 0 & 0 & 0 & 0 & 0\\
\hline
0 & 0 & \epsilon_s-t_{ss\sigma} & t_{sp\sigma} & 0 & 0 & 0 & 0\\
0 & 0 & t_{sp\sigma} & \epsilon_p-t_{pp\sigma} & 0 & 0 & 0 & 0\\
\hline
0 & 0 & 0 & 0 & \epsilon_p+t_{pp\pi} & 0 & 0 & 0\\
\hline
0 & 0 & 0 & 0 & 0 & \epsilon_p-t_{pp\pi} & 0 & 0\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & \epsilon_p+t_{pp\pi} & 0\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & \epsilon_p-t_{pp\pi}
\end{array}\right)
$$

Note that the Hamilton matrix here differs from the one given previously, because its is represented in a different basis, namely the symmetrized orbitals $|\phi_j\rangle$ as opposed to the original basis states $|\chi_\alpha\rangle$. The basis set plays the role of a coordinate system in Hilbert space. Like the components of a vector, also the matrix elements of a tensor depend on the choice of the coordinate system.

3.5.5 Diagonalize the sub-blocks

The subblocks can be diagonalized. We use the approximate expression Eq. 2.19 for the non-degerate two-center bond with $|t| < |\epsilon_1 - \epsilon_2|$. The modifications of the orbitals are demonstrated in Fig. 3.3 on p. 41.

- For the first block containing the *ssσ* and *ppσ* bonding orbitals we obtain the energies

$$
\begin{aligned}
\epsilon_- &\approx \epsilon_s + t_{ss\sigma} - \frac{|t_{sp\sigma}|^2}{\epsilon_p - \epsilon_s + t_{pp\sigma} - t_{ss\sigma}}\\
\epsilon_+ &\approx \epsilon_p + t_{pp\sigma} + \frac{|t_{sp\sigma}|^2}{\epsilon_p - \epsilon_s + t_{pp\sigma} - t_{ss\sigma}}
\end{aligned}
$$

  - The lower of the two states will have a character of a *ssσ* bond, however with a contribution from the *ppσ* bond mixed in such that the electrons are further localized in the bond.
  - The higher lying state will have a character of a *ppσ* bond, bit now the *ssσ* orbital is mixed in such that the electron density in the bond is depleted.

- For the second block containing the *ssσ* and *ppσ* antibonding orbitals we obtain the energies

$$
\begin{aligned}
\epsilon_- &\approx \epsilon_s - t_{ss\sigma} - \frac{|t_{sp\sigma}|^2}{\epsilon_p - \epsilon_s + t_{pp\sigma} - t_{ss\sigma}}\\
\epsilon_+ &\approx \epsilon_p - t_{pp\sigma} + \frac{|t_{sp\sigma}|^2}{\epsilon_p - \epsilon_s + t_{pp\sigma} - t_{ss\sigma}}
\end{aligned}
$$

- The lower of the two states will have a character of a $ss\sigma$ antibond, however with a contribution from the $pp\sigma$ antibond mixed in such that the electrons are less localized in the bond. The state becomes less antibonding, which lowers its energy.
- The higher lying state will have a character of a $pp\sigma$ antibond, bit the $ss\sigma$ orbital is mixed in such that the electron density in the bond is enhanced, making the state more antibonding, which shifts it up in energy.

The π type orbitals are already eigenstates, namely the bonding and antibonding states shown in Fig. 3.4.

3.6 Stability of main group dimers

From the qualitative arguments we would not be able to say if the σ-state in the middle is below or above the π states. From the node counting, we would expect the σ-state to lie above the *pi* states, because the latter have only one node plane, while the σ state has two. I believe that the reason is that a small perturbation such as an electric field perpendicular to the bond axis would deform the orbitals such that the two node planes would join and form one connected node plane.

The little diagram tells us already a lot about stability.

1. for alkali metal dimers such as Li_2, Na_2 etc. we expect a reasonably strong bond. The bonding is analogous to that in the hydrogen molecule.
2. On the other hand we expect Be_2 Mg_2 to be marginally stable, because there is one σ-bond and one σ^*-anti-bond, which largely cancel their effect. There will be some bonding though, because the σ-p orbitals mix into the lowest two states and thus stabilize them.
3. For the dimers B_2, Al_2, there will again be a net σ bond. Thus we expect its bonds to be clearly stronger than that of the dimers of the di-valent atoms.
4. Then we arrive at C_2 and Si_2 where the four lowest orbitals are occupied. Since we have two partially occupied degenerate orbitals, we need to be careful about spin-polarization. C_2 is a difficult molecule, because the ordering of π bonds and σ bond depends on the distance. When the σ bond is occupied, the system will have a net magnetic moment, because the two electrons can align their spins according to Hund's rule. When the σ orbital is unoccupied, the system will have a π double bond.
5. Now we come to the most stable ions such as N_2. Dinitrogen actually forms a triple bond. It consists of the two degenerate bonding π orbitals one one net σ bond. Dinitrogen is so non-reactive that it is often used to package food to avoid oxidation. Even though the atmosphere consists of 70% of nitrogen molecules, we need to supply nitrogen compounds in the form of fertilizer, to allow the excessive plant growth required for agriculture. Only some bacteria are able to transform dinitrogen to a form that can be metabolized.
6. Now we come to the chalcogenides[5] such as O_2 and S_2. Oxygen has two electrons in the π^* orbitals. Thus its net bonding corresponds to a double bond. However, it is stabilized by Hund's rule. The two electrons in the π^* states are spin paired. Thus O_2 has a net magnetic moment and a triplet spin splitting.That is if we apply a magnetic field, we will see three absorption lines.
7. The halogen dimers are again only weakly bond. They have a net σ bond, but all the gain from π bonding is compensated by the π^* anti-bonds.
8. The Nobel gases do not bind, because all bonding orbitals are compensated with anti-bonding orbitals.

---

[5] chalcogenide is spelled with a hard "ch" such as in "chemistry". Chalcogenides are oxides, sulfides, selenides, tellurides, that is compounds of the oxygen group. Oxides alone are not named chalcogenides unless grouped together with one of the heavier elements.

3.7 Bond types

- σ-bond. The wave function of a sigma bond is approximately axially symmetric. An example is the bond in the H_2 molecule.

- π-bond. A π bond has a node plane parallel to the bond axis. An example are two p orbitals which stand perpendicular to the bond axis. It is weaker than a σ bond because the orbitals do not point towards each other resulting in smaller matrix elements. An example is the planar molecule ethene (C_2H_2), where two p-orbitals on the carbon atom combine in a π bond in addition to the σ-bond

- **Double bond** is typically a π bond combined with a σ bond. The example is again ethene.

- A **triple bond** is two π-bonds combined with a single bond. An example is N_2.

- **Three-center bond** involves three orbitals in a row. A simple example is the hypothetical linear H_3 molecule. It forms a fully bonding orbital, a non-bonding orbital, to which the center atom does not contribute, and a fully anti-bonding orbital. The three center bond plays a role for five-valent atoms such as P. Such atoms form trigonal bi-pyramids. There are three equatorial sp^2 orbitals and a three-center bond along the z-axis.

- The best example for an **aromatic bond** is benzene (germ.:Benzol nicht Benzin!). It has the formula unit C_6H_6. The carbon form a six-fold ring. After we formed the sigma bonds there are six electrons available to form three double bonds. There are however not three double bonds and three single bonds but all the bond lengths are symmetric.

  The wave function is best constructed from symmetry arguments

  $$|\Psi_n\rangle = \sum_j |\chi_j\rangle e^{-2\pi i \frac{nj}{6}} \sqrt{\frac{1}{6}}$$

  where the χ_i are the p$_z$ orbitals on the individual carbon atoms. The Hamilton matrix has the form $H_{i,i} = \bar{\epsilon}$ and $H_{i,i+1} = H_{i-1,i} = -t$.

  The energies are $\epsilon_n = \bar{\epsilon} + t(e^{2\pi i n/6} + e^{-2\pi i n/6}) = \epsilon_n = \bar{\epsilon} - 2t\cos 2\pi n/6$. Thus there is the fully symmetric ground state at $\bar{\epsilon} - 2t$. Then there are two degenerate occupied states at $\bar{\epsilon} - t$, two degenerate unoccupied states at $\bar{\epsilon} + t$ and a fully anti-bonding state at $\bar{\epsilon} + t$

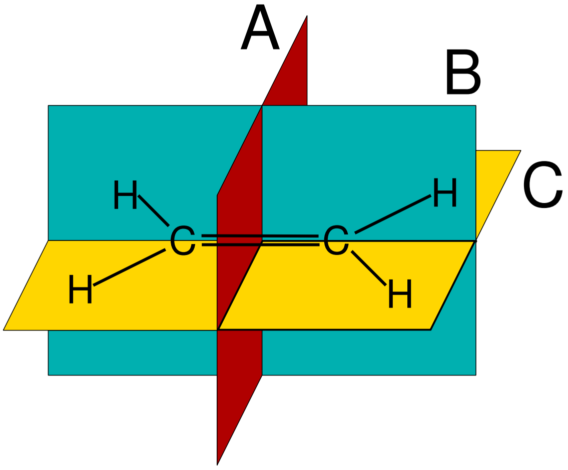
3.8 Worked example: orbitals of the ethene

The molecule **ethene**, also called ethylene, has the formula unit C_2H_4 and forms a planar structure, where each carbon atom has two bonds to hydrogen and one to the other carbon atom with a bond angle of approximately 120°.

Select Basisset

First we decide which orbitals we wish to include in the basis. We use the relevant orbitals, which are the four hydrogen s-orbitals, and the s- and p-orbitals on carbon. That is we work with twelve basis functions.

Determine symmetry operations

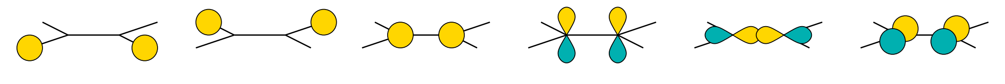
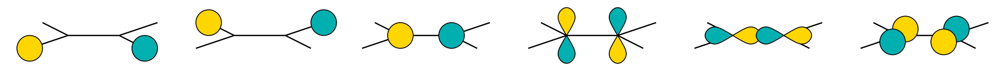


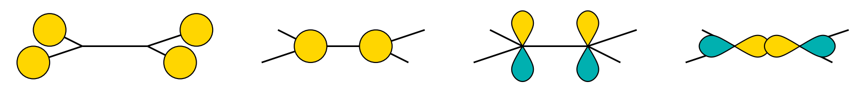
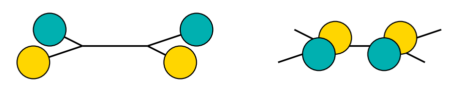
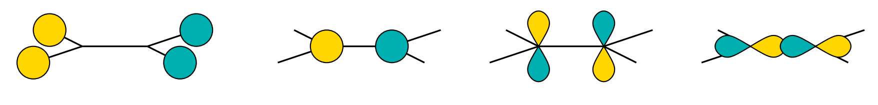
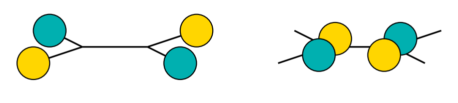
We begin determining the symmetry operators of methane.

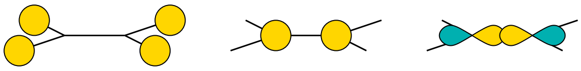
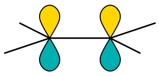
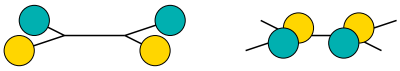
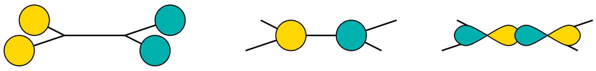
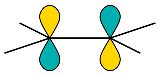
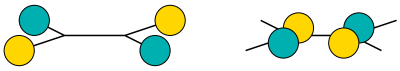
- A mirror plane in the molecular plane. (C)
- A mirror plane perpendicular to the C-C bond (A)
- a mirror plane containing the bond but perpendicular to the molecular plane (B)
- a two fold rotation about the bond axis

Determine symmetrized orbitals

- We start with the symmetry plane (A), perpendicular to the bond and determine the symmetric and antisymmetric orbitals.

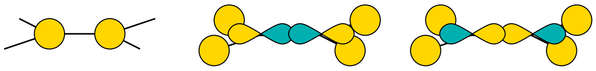
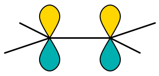
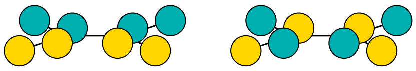
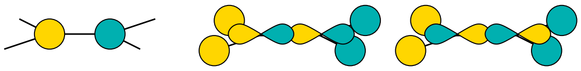
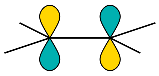
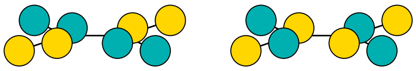
A+	
A–	

A+B+	
A+B–	
A–B+	
A–B–	

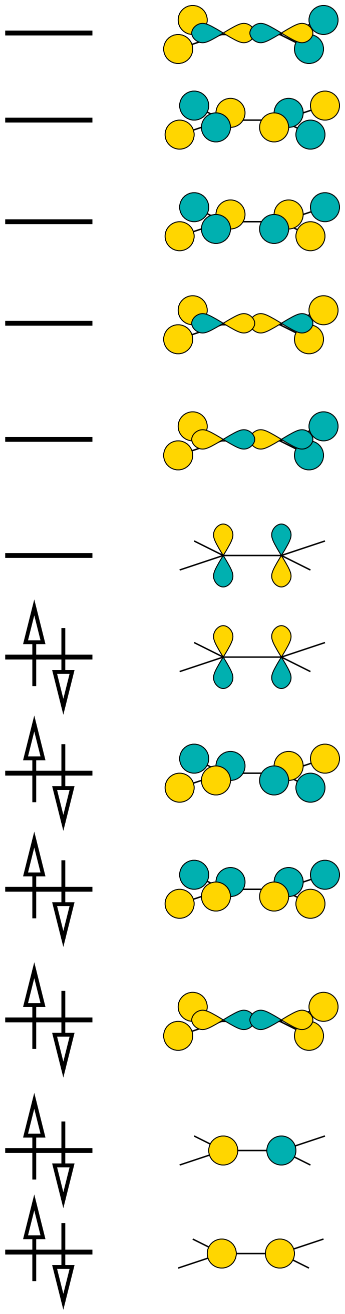
A+B+C+	
A+B+C−	
A+B−C+	
A−B+C+	
A−B−C−	
A−B−C+	

Form bond orbitals

We ended up with two 1 × 1 blocks, two 2 × 2 blocks and two 3 × 3 blocks of the Hamiltonian. Because the 3 × 3 blocks are difficult to handle, we form bonding and antibonding orbitals between the hydrogen orbitals and the carbon-p-orbitals, which are energetically in a similar position.

A+B+C+	
A+B+C−	
A+B−C+	
A−B+C+	
A−B−C−	
A−B−C+	

Now we can align the orbitals on an energy axis. Without a proper diagonalization this step can be nothing more than an educated guess. However it usually is sufficient to determine which orbitals are occupied and which are empty. Furthermore we can usually guess, which are the highest occupied and the lowest unoccupied orbitals, which is important for chemical considerations. We order the orbitals according to the bonding character. We place the s-orbitals on carbon lowest. Then we place the bonding orbitals low in energy, and the fully antibonding up. If there is a competition between σ and π-bonds, the σ bonds dominate.



Difficult is the decision for the two states above the lowest unoccupied state, because there is a competition between CH-bonds with CC-bonds. We could as well have placed them below the π antibond. These orbitals can hybridize a little with the carbon *s*-orbitals, which would place them higher up on energy, concentrating their weight onto the anti-bonds.

3.9 Worked example: orbitals of the methane

The molecule **methane** has the formula unit CH_3 and forms a tetrahedron. Let us work out its eigenstates:

Select Basisset

First we decide which orbitals we wish to include in the basis. We use the relevant orbitals, which are the four hydrogen s-orbitals, one carbon s-orbital and the three carbon p-orbitals. That is we work with eight basis functions.

Determine symmetry operations

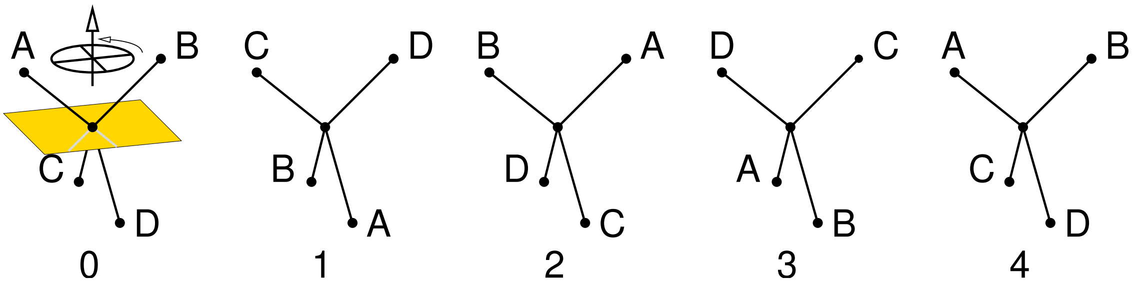
We begin determining the symmetry operators of methane.

- six mirror planes through a pair of hydrogen and the central carbon atom.
- four three fold axis along a C-H bond.
- six two-fold rotations about any axis through the central carbon and the mid-point between two hydrogen atoms.
- A rotation about a two-fold axis by 90°, followed by a mirror plane perpendicular to the axis passing through the central carbon atom.

Determine symmetrized orbitals

Now we need to select a number of symmetry operators which commutate.. Two transformations commutate, if we transform an object in sequence by both transformations, and the result is identical irrespective of the object chosen. For example two reflections with perpendicular mirror planes commutate. They do not commutate if they have another angle.

Let us start with the rotation by 90° followed by a reflection. This operation brings an object into its original position after it has been applied four times.



This operation produces the identity after four-fold application. The eigenvalues are therefore $1, i, -1, -i$.

- Let us construct all orbitals for the eigenvalue 1. according to Eq. 3.10.
    - We start with one hydrogen orbital on the site A. We transform it once by the symmetry operation and multiply with the eigenvalue, namely 1, which results in a hydrogen orbital with the same sign on site D. We take this orbital, apply the symmetry operation again and obtain, after multiplication with 1, an orbital at site B . The third operation produces an orbital at site B. We add the four orbitals together, and obtain the orbital shown in Fig. 3.7.

      We see immediately, that we obtain the same orbital, irrespective of which orbital we started out from.

  - Now we proceed with the s-orbital on the carbon atom. This orbital falls on itself after each transformation. Since its copies are multiplied with 1, the orbital remains unchanged.
  - Next we proceed to the *p*-orbitals. Because the p_z orbital changes its sign after each transformation, its copies add to zero. The p_x and p_y orbitals change their sign after every two operations and therefore also cancel out.

  In total we obtain exactly two symmetrized orbital with eigenvalue 1 of the symmetry operation.

- We proceed to eigenvalue -1.
  - We begin again with a hydrogen orbital. After each transformation we need to multiply the orbital with -1. Thus the orbital enters with a positive sign on sites A and B and with negative sign on sites C and D. After adding up we obtain the symmetrized orbital shown in Fig. 3.7 in the box with eigenvalue -1.
  - the carbon *s* orbital does not produce a symmetrized orbital in this class, because its copy add up to zero after multiplication with the factors 1 and -1.
  - The p_z orbital is a symmetry orbital in this class, because the symmetry operation simply reverts its sign. After multiplication with the eigenvalue, namely -1, we obtain the original orbital.
  - the p_x and p_y orbitals are rotated by 90° in each operation. That is, after two operations, they revert their sign. After two operations we need to multiply with the square of the eigenvalue. Thus the copies of the orbitals cancel each other, and do not contribute to this symmetry class.

  Also in this class we obtain exactly two orbitals.

- Now we proceed to the complex valued eigenvalue i.
  - We start with an *s* orbital and obtain two real-valued orbitals with opposite sign on the upper two atoms, and two imaginary-valued orbitals with opposite sign on the lower two hydrogen atoms. Thus the symmetrized orbital is complex valued.
  - the s-orbital on carbon cancels out, because it falls on itself unter the operation and $i + i^2 + i^3 + i^4 = i - 1 - i + 1 = 0$.
  - Also the p_z orbital cancels out, because it changes it sign and $i - i^2 + i^3 - i^4 = 0$.
  - the p_x and p_y orbitals are transformed into each other by the symmetry operation. From p_x we obtain a symmetrized orbital $\frac{1}{\sqrt{2}}\left(|p_x\rangle + i|p_y\rangle\right)$. The p_y orbital produces the same symmetrized orbital with an additional factor i. Note, that this orbital is en eigenstate to $\hat{L}_z$ with eigenvalue $m = 1$.
- for the eigenvalue $-i$, we obtain the same symmetrized orbitals as for the eigenvalue i, only that the orbitals are complex conjugate.

Thus we have divided the 8×8 Hamiltonian in four 2×2 blocks.

Form bond and antibonding orbitals

Next we diagonalize the 2×2 blocks of the Hamiltonian expressed by symmetrized orbitals and thus obtain bonding and antibonding orbitals.

In order to align the orbitals on an energy axis we need to know the energies and hopping matrix elements. Because the carbon s-orbital lies energetically below the carbon p-orbital, we can conclude that the average value of the bonding and antibonding orbitals that contain the carbon s-orbital lies below that of the p-orbitals. Furthermore we recognize that the type and number of interactions for all bonding orbitals containing the carbon p-orbital is the same. Thus we conclude that they are degenerate. The same holds true for the corresponding antibonding orbitals. This leads us to the schematic assignment of the electron states of methane shown in Fig. 3.9.

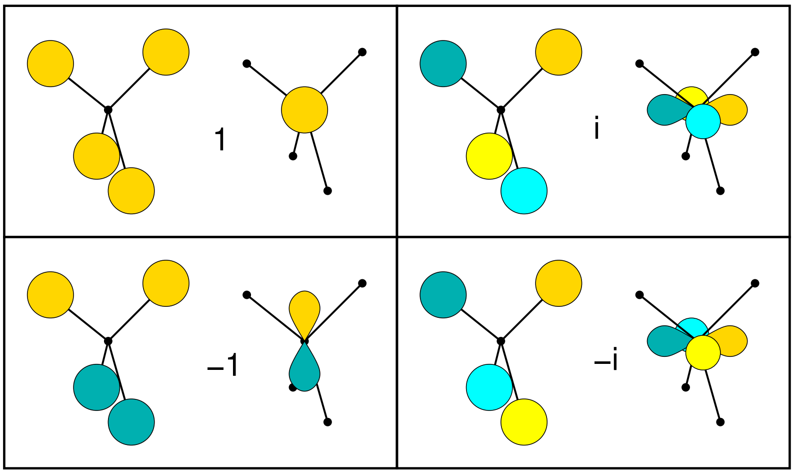


Fig. 3.7: Symmetrized orbitals of methane. The numbers refer to the eigenvalue of the symmetry operation, namely a fourfold rotation about a vertical axis combined with a mirror operation about the horizontal mid-plane. The golden and green lobes have values $+1$ and -1, and the bright yellow and cyan lobes have values $+i$ and $-i$ respectively. .

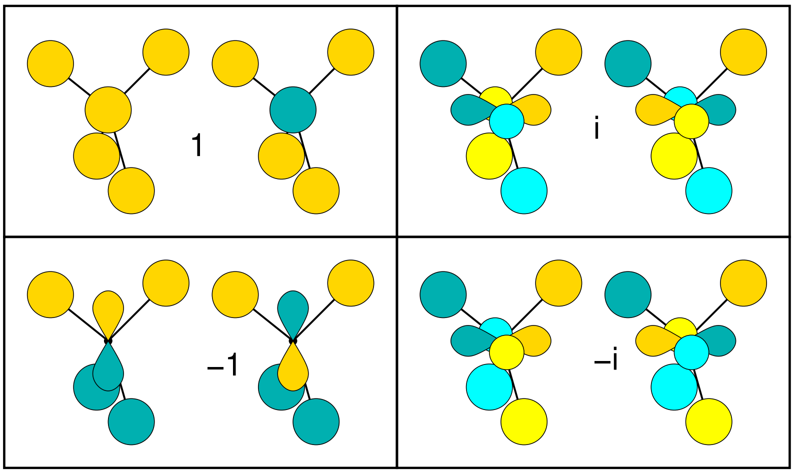
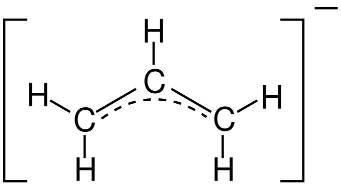


Fig. 3.8: Bonding and antibonding orbital, according to symmetry class. .

3.10 Worked example: orbitals of the allyl ion

Let us work out the molecular orbitals of the allyl anion $C_3H_5^-$.



The allyl ion is a planar molecule.

- First we work out the symmetry operations of the molecule. The molecule has the following

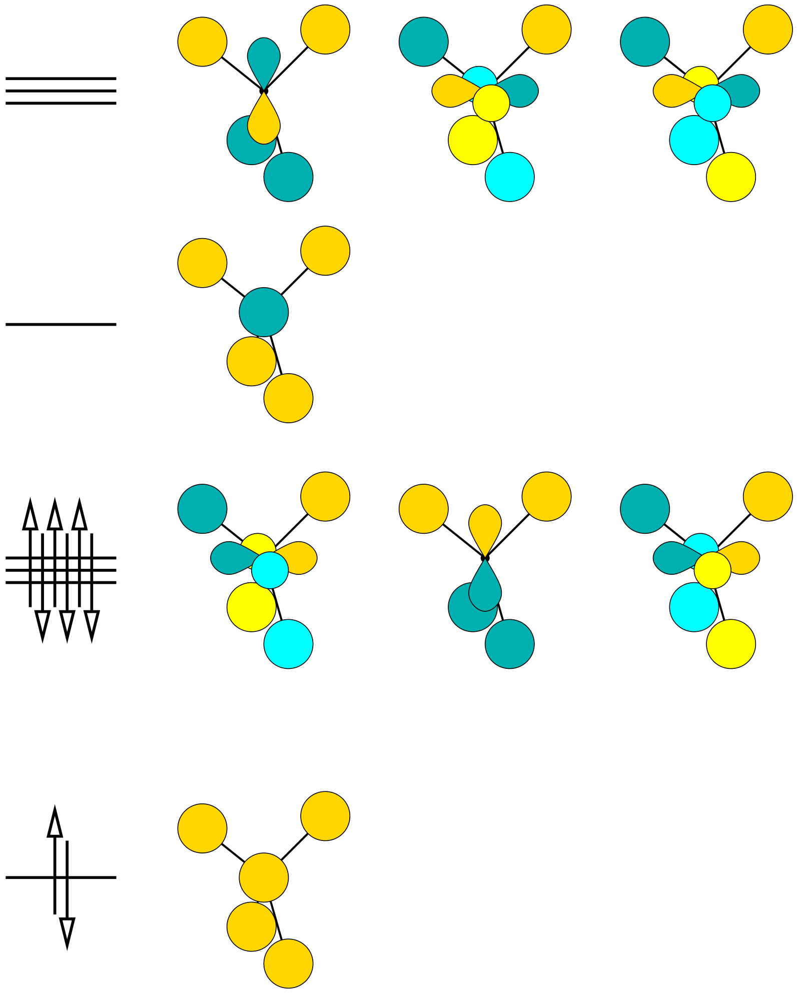
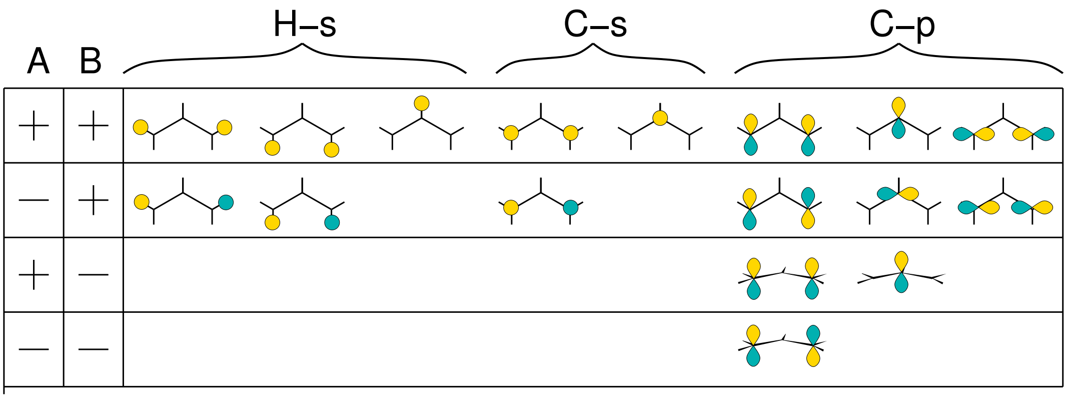


Fig. 3.9: Symmetry eigenstates and energy level ordering of methane. The blue and bright-yellow lobes correspond to the imaginary part of the wave function. The imaginary part can be avoided by a suitable superposition of the two wave functions that are complex conjugate of each other. While the resulting states are still eigenstates of the Hamilton operator, they are no more eigenstates of the symmetry operators. .

symmetry operations.

- a mirror plane in the plane of the molecule
- a mirror plane perpendicular to the molecular plane, containing the central C-H bond.
- a two-fold rotation about an axis through the central C-H bond.

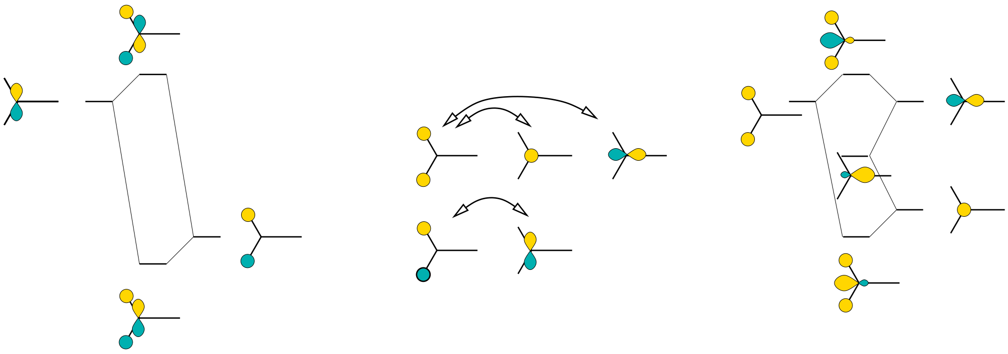
- Now we work out the symmetrized orbitals. We start with every "atomic" orbital and construct the symmetrized orbital according to Eq. 3.10 by adding the orbitals obtained from the first by applying the symmetry operation repeatedly and multiplication with the eigenvalue of the symmetry operation to the corresponding power.



- We check the number of symmetrized orbitals, which must be identical to the number of orbitals we started from: There have been 5 hydrogen s-orbitals, 3 carbon s-orbitals and 9 carbon p-orbitals, that is in total 17 orbitals. We also have 17 symmetrized orbitals.

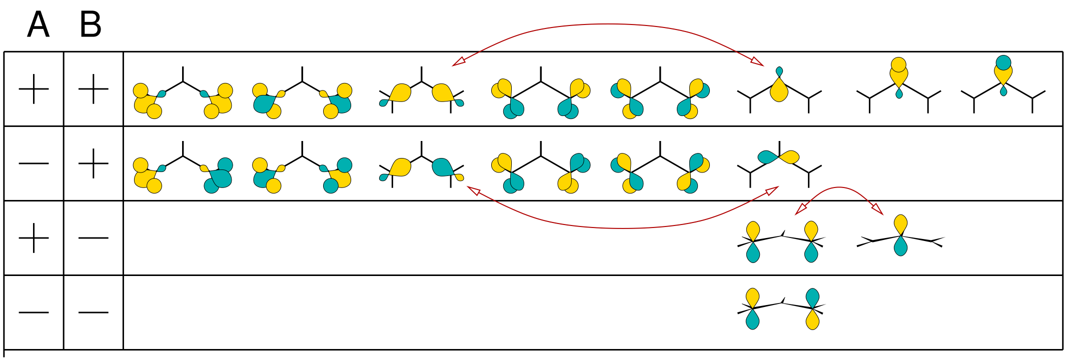
- Since we ended up with sub blocks of the Hamiltonian of up to eight orbitals we need to go a step further. We use an approximate symmetry of the CH_2 end groups and determine symmetrized orbitals for these fragments. We are careful not to mix states from different symmetry classes of the entire molecule, in order to avoid destroying the block-diagonalization of the Hamiltonian reached so far.

  – We determine the symmetry groups of the end group.

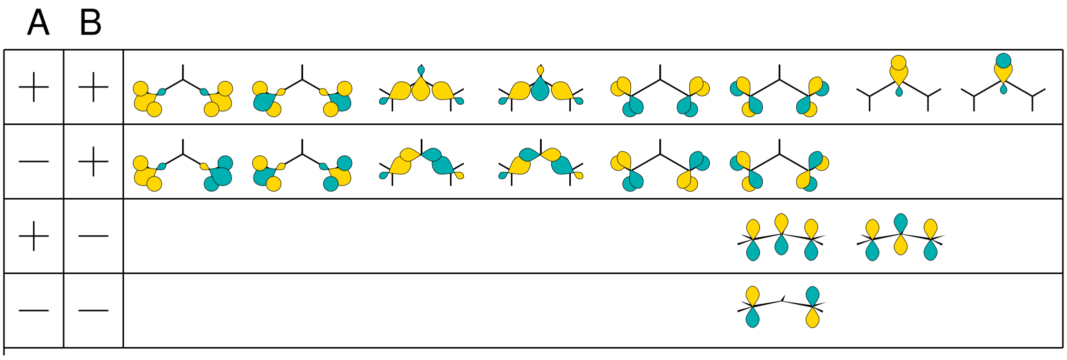


  –

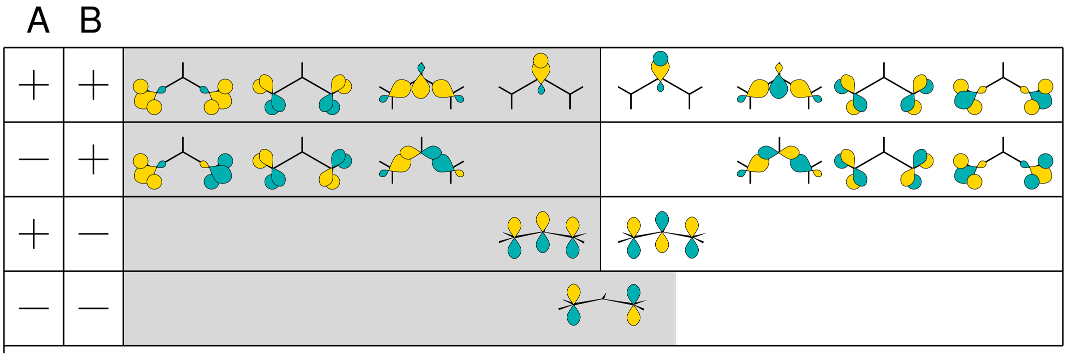
- now we compose again the symmetrized orbitals with the locally symmetrized end groups



- We notice that there are only few orbitals that still interact with each other, if we ignore the weak interactions of orbitals pointing away of each other. We diagonalize the last few 2×2 matrices and thus obtain an approximate idea of the eigenstates.



- We divide the orbitals into bonding and antibonding orbitals. This is not necessarily as straightforward as in this case. Then we count the number of valence electrons, which is 18, resulting in 9 filled orbitals.We fill up the bonding orbitals first, proceeding to the non-bonding orbitals. We find that the non-bonding π orbital is the highest occupied orbital.



The wave functions calculated from first principles are shown below in figure 3.10. We recognize good agreement for some of the wave functions and poor agreement for others. We have been able to identify the highest occupied orbital. We also recognize that the main character of the states is similar in our estimated orbitals and the ab-initio wave functions.

The main error is that states lying in a similar energy region are hybridized incorrectly: That is a superposition of those states from our estimate would allow us to construct the true eigenstates. The limitations of the above procedure are not untypical. Like in this case, however, they usually do not strongly affect stability considerations. The reason behind this is that a unitary transformation of completely filled or of completely empty states does not affect the energy. If the intermixing between filled and occupied orbitals would be equally undetermined, the results would not be reliable, and one should resort to ab-initio calculations.

3.11 Worked example: orbitals of closo-$B_6H_6^{2-}$

Boron can form B_6^{2-} octahedra, which can arrange to crystals in a simple cubic lattice to form hexaborides. In the form of $B_6H_6^{2-}$ the boron cluster exists as anion in salts.

The discussion of bonding[**?**] of the hexaborides starts from tangential and radial orbitals. The radial orbitals are the sp-hybrids[6] pointing inward and outward of the cage. The tangential orbitals are p-orbitals perpendicular to the direction to the cage-center.

The symmetry-adapted orbitals of the hexaboride cage are shown in figure 3.11. The symmetry operations considered are the mirror reflections on the three planes, which cut through the center and four opposite corners of the octahedron.

---

[6] Hybrid orbitals are discussed in te next chapter.

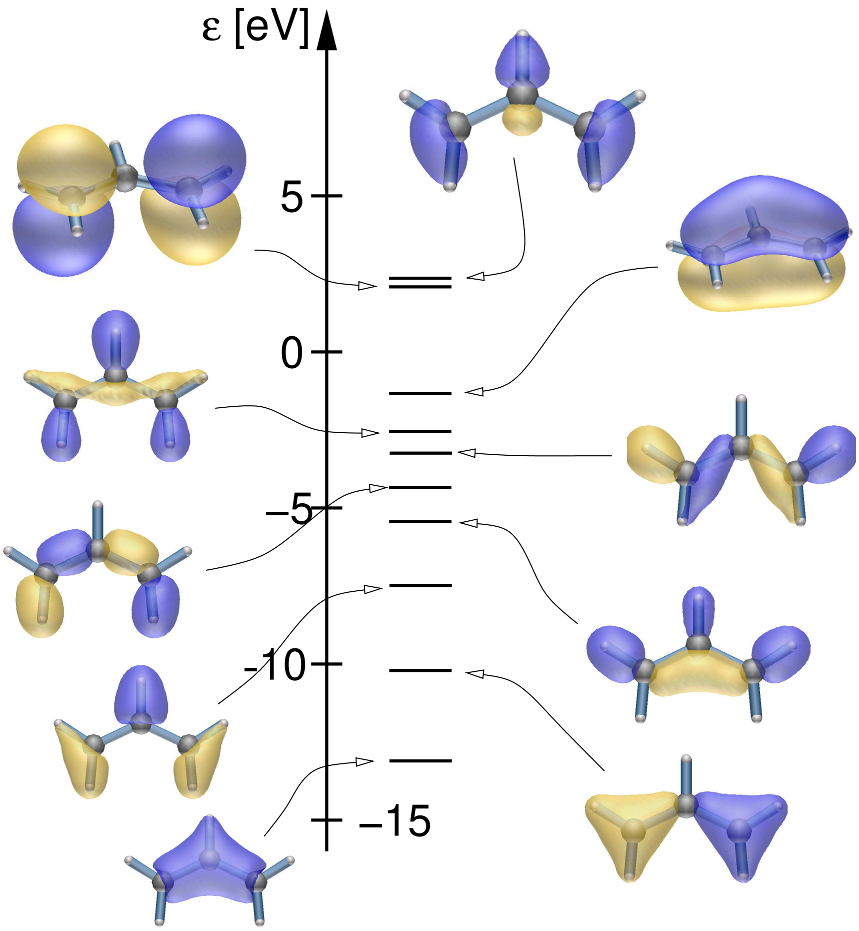


Fig. 3.10: PAW calculation of the one-particle orbitals of the allyl ion. The highest state shown is mostly localized in the vacuum.

- The orbitals in the botton line are symmetric under all mirror operations are symmnetric under all three mirror reflections. The second and the third orbitals are analogous to the two e_g orbitals of a d-shell. Analogous to the sp-hybrids pointing inward, also the dangling bonds can also be combined into a fully symmetric combination on the one side and a pair of e_g-like combinations.
- The left box in the three top rows are symmetric with respect to one mirror operation and antisymmetric with respect to the remaining two. The two orbitals in each box do not have a Hamilton matrix element connecting them, because they transform differently with respect to another mirror re3flection.
- The central and right boxes in the three top rows are symmetric with respect to two mirror operation and antisymmetric with respect to the remaining two. The central and right boxes have distinct eigenvalues (1 repectively -1) with respect to the four-fold rotation about the special axis.

The construction of the Hamilton eigenstates is not unique. The high symmetry of the octahedron produces singlets, doublets and triplets. The choice or orbitals in the degenerate multiplets are arbitrary.

One choice of eigenstates of the Hamiltonian of the hexaboride cage are described below. The order of states and the symmetry labels are taken from schmitt et al.[12].

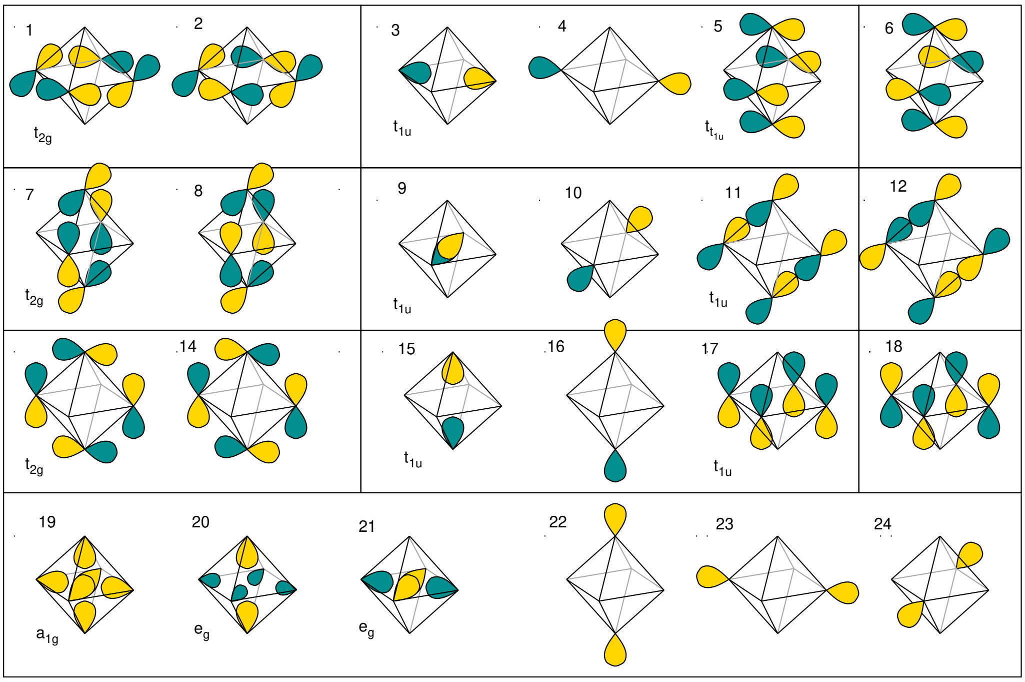


Fig. 3.11: Symmetry-adapted orbitals of the hexaboride cage. For explanation, see text. Editor: The symmetry labels need to be confirmed.

1. $a1g$: Of the orbitals in figure 3.11, the state Nr. 19 at the bottom left corner will have the lowest energy. It is a purely symmetric combinations of sp-hybrids pointing into the center of the cage.
2. t_{1u}: triplet (orbitals Nr. 5, 11, 17) of equatorial p-*pi*-bond-ring in a bonding combination with the axial sp-orbitals pointing inward (orbitals Nr. 3, 9, 15). (left-most and the right-most orbitals of the central boxes of the top three rows.)
3. t_{2g}: The equatorial p-σ-bond-rings form a triplet. (Left orbital of the top three rows). Each bond has an admixture of a p-*pi* bond.
4. The six dangling bonds at the corner (orbitals 22,23,24,4,10,16) consume six electrons to form bonds with their neighbors. The dangling bond states form one singlet, one doublet and one triplet.
5. t_{2u}? This triplet forms the LUMO of $B_6H_6^{2-}$ and it consists of π^* antibonds (orbitals Nr. 6, 12, 18).
6. t_{1g}?: the antibonds between the equatorial *pi*-bonds with the axial sp-hybrids pointing inward form the second-highest triplet.
7. e_g: A doublet of two states which are symmetric under all three mirror reflections considered. (orbitals Nr. 20, 21).
8. t'_{1u}?: is the triplet of the the p-σ^* antibonds along a square. (orbitals Nr. 2, 8, 14).

Editor: I have not verified the symmetry labels with the question marks.

For an electroni9c structure calculation, see Schmitt et al.[12].

For the classification and the electron count of boron hydrides see Wade's rules.

3.12 Exercise: sketch orbitals

- sketch wave functions and energy level diagram for a water molecule. The water molecule H_2O has a bond angle of about 104°. The hydrogen orbitals are located slightly above the oxygen p orbitals.
- The cyclopentadienyl ion has the formula $[C_5H_5]^-$. It is a five-membered ring. Work out first the orbitals for one bead and use then the knowledge about eigenstates for ring systems.
- sketch the wave functions of ethene (ethylene) C_2H_4.

Chapter 4

Hybrid orbitals

Sofar we have used only two orbitals at a time that form a bond. In molecules there are many more electrons. In order to arrive even approximately at a description of a bond by two orbitals, we need to design so-called hybrid orbitals. Hybrid orbitals are constructed from atomic orbitals in such a way that they point optimally towards the neighbors.[1]

4.1 Spherical and cubic harmonics

The spherical symmetry of the atom implies conservation of the angular momentum. Thus it is no surprise that the eigenstates of the angular momentum are important to describe the atomic wave functions.

The eigenstates of the angular momentum are the **spherical harmonics**. Because the different components of the angular momentum operator $\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}}$ do not commutate, the two relevant eigenvalue equations are those for the total angular momentum or better for $\hat{\vec{L}}^2$ and one component that usually is chosen as the z-component $\hat{L}_z$.

The eigenvalue equations for the spherical harmonics are

$$\begin{aligned}
\hat{\vec{L}}^2|Y_{\ell,m}\rangle &= |Y_{\ell,m}\rangle\hbar^2\ell(\ell+1) \qquad && \text{with } \ell = 0, 1, 2, \ldots \\
\hat{L}_z|Y_{\ell,m}\rangle &= |Y_{\ell,m}\rangle\hbar m && \text{with } m = -\ell, -\ell+1, \ldots, \ell-1, \ell
\end{aligned}$$

ℓ is called the **angular-momentum quantum number**. The atomic orbitals are usually referred to as s, p, d, f-orbitals, if they have the angular momentum quantum number $\ell = 0, 1, 2, 3$, respectively. m is the **magnetic quantum number** , because the m-degeneracy can be lifted by applying a magnetic field.

Spherical harmonics $Y_{\ell,m}(\vec{r})$ only depend on the direction of $\vec{r}$ and are independent of the radius $|\vec{r}|$. This is the reason for expressing them in polar coordinates as function of the angular coordinates θ and ϕ.

The definition of spherical harmonics is not unique. They differ in the so-called Condon-Shortley phase $(-1)^m$ and the definition of the normalization.

We use the normalization

$$\frac{1}{r_c^2}\oint_{|\vec{r}|=r_c} dA\, Y^*_{\ell,m}(\vec{r})Y_{\ell',m'}(\vec{r}) = \delta_{\ell,\ell'}\delta_{m,m'} \tag{4.1}$$

---

[1] Here we consider only hybrid orbitals for s- and p orbitals. There are similar constructions including d-orbitals. The latter are less useful, since the d-like valence electrons behave so different from the s- and p orbitals, that combining them into hybrids is rarely useful.

In polar coordinates[2] the spherical harmonics have the form

$$Y_{\ell,m}(\theta,\phi)=\sqrt{\frac{2\ell+1}{4\pi}\cdot\frac{(\ell+m)!}{(\ell-m)!}}\quad\underbrace{P_\ell^m(\cos(\theta))}_{\substack{\text{associated}\\\text{Legendre polynomial}}}\quad \mathrm{e}^{im\phi} \tag{4.2}$$

$$Y_{\ell,-m}(\theta,\phi)=(-1)^m\,Y^*_{\ell,m}(\theta,\phi) \tag{4.3}$$

Here $\cos(\theta)=z/|r|$ and $\cos(\phi)=\frac{x}{\sqrt{x^2+y^2}}$ $\sin(\phi)=\frac{y}{\sqrt{x^2+y^2}}$.

Cubic spherical harmonics

In addition to the regular spherical harmonics there are the cubic or **real spherical harmonics** $\bar{Y}_{\ell,m}(\vec{r})$

$$\bar{Y}_{\ell,m}=\sqrt{2}\,\mathrm{Re}[Y_{\ell,|m|}]$$
$$\bar{Y}_{\ell,-m}=\sqrt{2}\,\mathrm{Im}[Y_{\ell,|m|}]$$

The real spherical harmonics are still eigenstates of $\hat{\vec{L}}^2$, but **they are no eigenstates of $\hat{\vec{L}}_z$**

The main reason for introducing real spherical harmonics is that can draw them easily, because they do not have an imaginary part.

Below, I list the most important spherical harmonics.

ℓ	m	(eigenstate of $\hat{L}_z$) $Y_{\ell,m}(r)$	(real) $\bar{Y}_{\ell,m}$
0	0	$\frac{1}{\sqrt{4\pi}}$	$\frac{1}{\sqrt{4\pi}}$
1	-1	$\sqrt{\frac{3}{4\pi}}\frac{x+iy}{\|\vec{r}\|}$	$\sqrt{\frac{3}{4\pi}}\frac{x}{\|\vec{r}\|}$
1	0	$\sqrt{\frac{3}{4\pi}}\frac{z}{\|\vec{r}\|}$	$\sqrt{\frac{3}{4\pi}}\frac{z}{\|\vec{r}\|}$
1	1	$\sqrt{\frac{3}{4\pi}}\frac{(x-iy)}{\|\vec{r}\|}$	$\sqrt{\frac{3}{4\pi}}\frac{y}{\|\vec{r}\|}$
2	-2	$\sqrt{\frac{15}{32\pi}}\frac{x^2-y^2+2ixy}{\vec{r}^2}$	$\sqrt{\frac{15}{32\pi}}\frac{x^2-y^2}{\vec{r}^2}$
2	-1	$\sqrt{\frac{15}{8\pi}}\frac{xz+iyz}{\vec{r}^2}$	$\sqrt{\frac{15}{8\pi}}\frac{xz}{\vec{r}^2}$
2	0	$\sqrt{\frac{5}{16\pi}}\frac{3z^2-r^2}{\vec{r}^2}$	$\sqrt{\frac{5}{16\pi}}\frac{3z^2-r^2}{\vec{r}^2}$
2	+1	$\sqrt{\frac{15}{8\pi}}\frac{xz-iyz}{\vec{r}^2}$	$\sqrt{\frac{15}{8\pi}}\frac{yz}{\vec{r}^2}$
2	+2	$\sqrt{\frac{15}{32\pi}}\frac{x^2+y^2-2ixy}{\vec{r}^2}$	$\sqrt{\frac{15}{8\pi}}\frac{xy}{\vec{r}^2}$

Graphical representation of real spherical harmonics

Spherical harmonics are usually represented by spheres and objects with different number of lobes. The surface of these objects are defined by $\frac{1}{|r|}Y_{\ell,m}(r)=\pm1$. The distance from the origin is the magnitude of the spherical harmonic.

Note, however, that one can also use any other radial part instead of $\frac{1}{|\vec{r}|}$, which has a small influence on the shape of the objects. Thus instead of two spheres, one may observe lobes with slightly different shapes.

---

[2] We use the polar coordinates defined by $x=rcos(\phi)\sin(\theta); y=r\sin(\phi)\sin(\theta); z=r\cos(\theta)$.

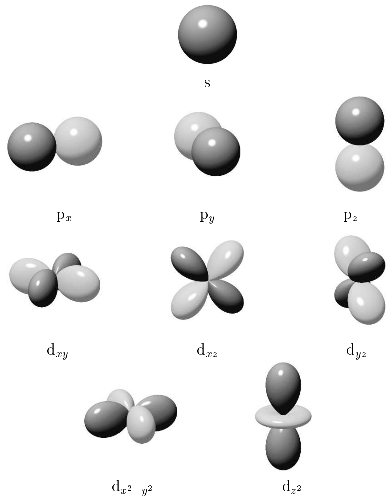


Fig. 4.1: Real spherical harmonics: the representation is a surface defined by $|\frac{1}{|r|}\bar{Y}_{\ell,m}(\vec{r})| = 1$. For a given direction, the distance of the surface from the origin is the absolute value of the cubic spherical harmonics. The different values of grey reflect to the signs of the function.

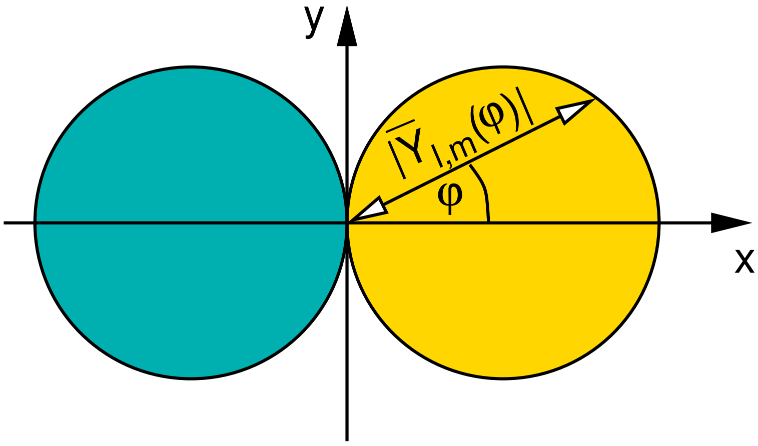


Fig. 4.2: Construction of the graph of a real spherical harmonics ($\cos(\varphi)$) in polar coordinates . Note that real and imaginary part need to be represented individually.

4.2 Rotation of p-orbitals

If we superimpose different p-orbitals we obtain other p-orbitals. A p-orbital has the form

$$\chi_{p_x}(\vec{r}) = R_p(|\vec{r}|)\sqrt{\frac{3}{4\pi}}\frac{x}{|\vec{r}|}$$
$$\chi_{p_y}(\vec{r}) = R_p(|\vec{r}|)\sqrt{\frac{3}{4\pi}}\frac{y}{|\vec{r}|}$$
$$\chi_{p_y}(\vec{r}) = R_p(|\vec{r}|)\sqrt{\frac{3}{4\pi}}\frac{z}{|\vec{r}|}$$

where $R_p(r)$ is the radial part of the orbital and Y_{p_x}, Y_{p_y} and Y_{p_z} are the real spherical harmonics for $\ell = 1$.

Thus if we wish to construct a p orbital that looks into the direction $\vec{e}$, where $\vec{e}$ is any normalized vector, we simply need to superimpose the three p-orbitals according to

P-ORBITAL WITH SPECIFIED ORIENTATION

For a normalized direction vector $\vec{e}$, a $p-$type orbital pointing in the direction of $\vec{e}$ is obtained by superposing the p_x, p_y, p_z orbtials as follows:

$$\chi_{p,\vec{e}}(\vec{r}) = \sum_{i=1}^{3} e_i \chi_{p,i}(\vec{r}) = R_p(|\vec{r}|)\sqrt{\frac{3}{4\pi}}\frac{\vec{e}\vec{r}}{|\vec{r}|}$$

$$|p_{\vec{e}}\rangle = |p_x\rangle e_x + |p_y\rangle e_y + |p_z\rangle e_z \tag{4.4}$$

where $R(r)$ is the radial part of the orbial $\chi(r)$.

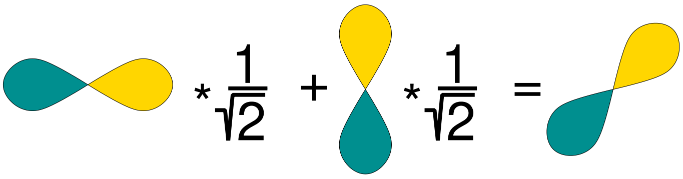


Fig. 4.3: Schematic representation of mixing p-orbitals among each other. The result is a rotated p-orbital

Thus, the shape of p orbitals does not change if we superimpose them, but only their direction.

4.3 Superposition of s- and p-orbitals: Hybrid orbitals

Similarly to the superposition of p-orbitals, we may investigate superpositions of s and p-orbitals.

$$\chi_{sp_x,\phi}(\vec{r}) = \chi_s(\vec{r})\cos(\phi) + \chi_{p_x}(\vec{r})\sin(\phi)$$

The angle ϕ is not an angle in the common sense: I have chosen the factors $\cos(\phi)$ and $\sin(\phi)$ only to ensure that the resulting orbital is again normalized.[3]

The resulting orbital has the form

$$\chi_{sp_x,\phi}(\vec{r}) = R_s(|\vec{r}|)\sqrt{\frac{1}{4\pi}}\cos(\phi) + R_p(|\vec{r}|)\sqrt{\frac{3}{4\pi}}\frac{x}{|\vec{r}|}\cos(\phi)$$

As we increase ϕ from zero to 90° the s-orbital is converted into a p-orbital.

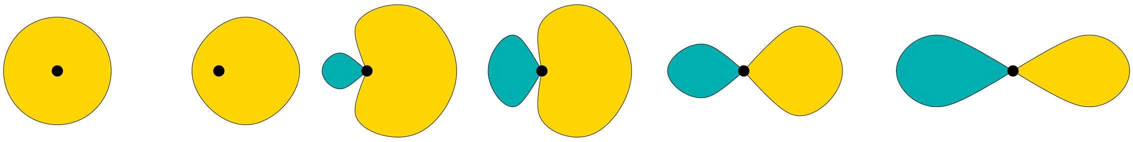


Fig. 4.4: Schematic representation of mixing s and p-orbitals. The series begins on the left with a pure s-orbital. As we proceed towards the right, an increasing amount of p-character is included, until one arrives at a pure p-orbital on the very right. The orbitals between the extremes are s-p hybrids.

The concept of hybrid functionals is very useful for hybrid orbitals between $s-$ and $p-$orbitals, because their wave functions have a very similar radial dependence. Often the concept is extended to $d-$ and $f-$ electrons. Personally I do not consider this very helpful for the following reason: while the energies of $d-$ and $f-$ states lie in the valence region, their shape is very core-like. That is, they barely extend beyond the atom core. Hybrid orbitals orbitals between core-like and extended orbitals can no more be described well by a new angular behavior and the nature of the orbital depends strongly on the distance. This renders the concept of $s-d$-like orbitals less useful.

---

[3] Normalization of the new orbital follows from $\sin^2(\phi)+\cos^2(\phi)=1$ and the orthogonality of the s- and p-orbitals.

The hybrid orbitals that can be formed from $s-$ and $p-$orbitals are shown in Fig. 4.5. In principle, we can form any mixture of $s-$ and $p-$orbitals. However, there are only a certain number of orthogonal sets of hybrid orbitals with pure p-orbitals. Those are

- unhybridized orbitals
- sp-hybrids
- sp^2 hybrids and
- sp^3 hybrids and

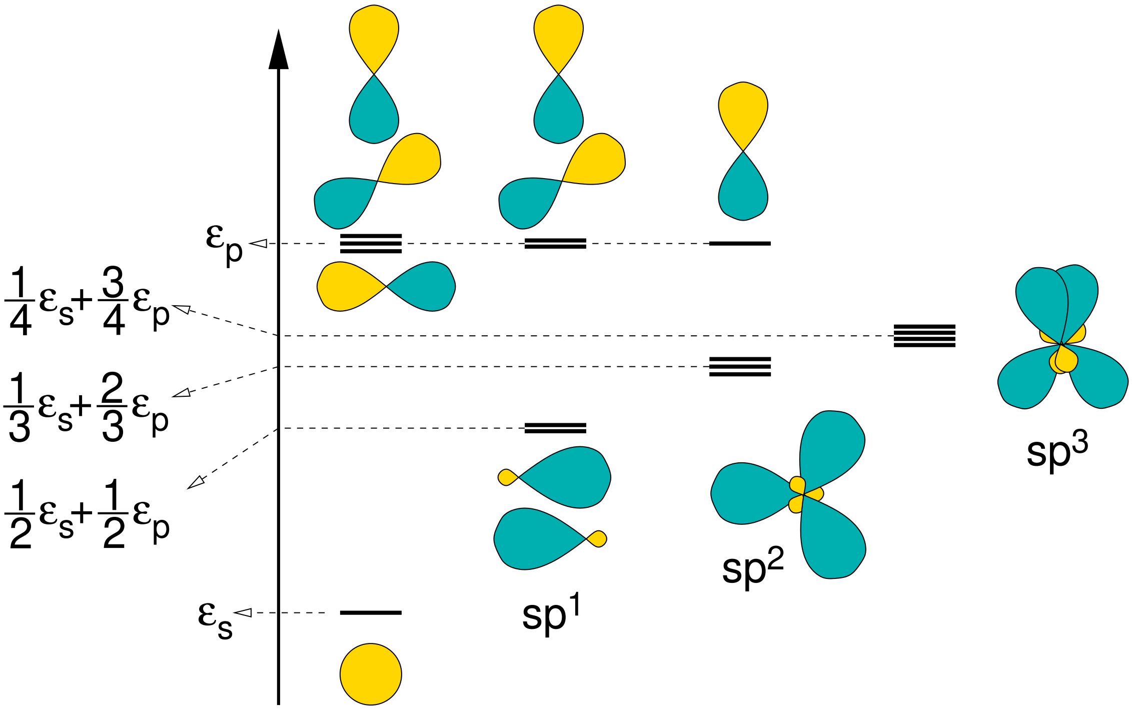

ϵ_p
$\frac{1}{4}\epsilon_s+\frac{3}{4}\epsilon_p$
$\frac{1}{3}\epsilon_s+\frac{2}{3}\epsilon_p$
$\frac{1}{2}\epsilon_s+\frac{1}{2}\epsilon_p$
ϵ_s
sp^1
sp^2
sp^3

Fig. 4.5: s- and p-orbitals can be mixed into hybrid orbitals that still form a orthonormal basisset.

- sp orbitals are used for linear molecules. One chooses a particular axis as the z-axis.

$$|sp_+\rangle = \frac{1}{\sqrt{2}}\Big(|s\rangle + |p_z\rangle\Big)$$
$$|sp_-\rangle = \frac{1}{\sqrt{2}}\Big(|s\rangle - |p_z\rangle\Big)$$
$$|p_x\rangle = |p_x\rangle$$
$$|p_y\rangle = |p_y\rangle$$

- sp^2 orbitals are important for planar materials such as graphite. There are three hybrid orbitals in a plane and one p orbital perpendicular to the plane. The hybrid orbitals have equal angles

(120^o) among each other and have the same Hamilton matrix element $\bar\epsilon = \frac{1}{3}\epsilon_s + \frac{2}{3}\epsilon_p$.

$$
\begin{aligned}
|sp^2_1\rangle &= |s\rangle\sqrt{\frac{1}{3}} + |p_x\rangle\sqrt{\frac{2}{3}} \\
|sp^2_2\rangle &= |s\rangle\sqrt{\frac{1}{3}} - |p_x\rangle\sqrt{\frac{1}{6}} + |p_y\rangle\sqrt{\frac{1}{2}} \\
|sp^2_3\rangle &= |s\rangle\sqrt{\frac{1}{3}} - |p_x\rangle\sqrt{\frac{1}{6}} - |p_y\rangle\sqrt{\frac{1}{2}} \\
|p_z\rangle &= |p_z\rangle
\end{aligned}
$$

- sp^3 orbitals are the structural motif of most organic materials and of solids such as diamond. There are four hybrid orbitals pointing into the corners of a tetrahedron. All have the same Hamilton matrix element $\bar\epsilon = \frac{1}{4}\epsilon_s + \frac{3}{4}\epsilon_p$.

$$
\begin{aligned}
|sp^3_1\rangle &= \frac{1}{2}\Big(|s\rangle + |p_x\rangle + |p_y\rangle + |p_x\rangle\Big) \\
|sp^3_2\rangle &= \frac{1}{2}\Big(|s\rangle - |p_x\rangle - |p_y\rangle + |p_x\rangle\Big) \\
|sp^3_3\rangle &= \frac{1}{2}\Big(|s\rangle + |p_x\rangle - |p_y\rangle - |p_x\rangle\Big) \\
|sp^3_4\rangle &= \frac{1}{2}\Big(|s\rangle - |p_x\rangle + |p_y\rangle - |p_x\rangle\Big)
\end{aligned}
$$

It may be of interest to compare the hybrid orbitals if they point into the same direction.

$$
\begin{aligned}
|s\rangle &= |s\rangle\frac{1}{\sqrt{1}} + |p_z\rangle\sqrt{\frac{0}{1}} = |s\rangle\cos(0^\circ) + |p_z\rangle\sin(0^\circ) \\
|sp^1\rangle &= |s\rangle\frac{1}{\sqrt{2}} + |p_z\rangle\sqrt{\frac{1}{2}} = |s\rangle\cos(45^\circ) + |p_z\rangle\sin(45^\circ) \\
|sp^2\rangle &= |s\rangle\frac{1}{\sqrt{3}} + |p_z\rangle\sqrt{\frac{2}{3}} = |s\rangle\cos(54.73^\circ) + |p_z\rangle\sin(54.73^\circ) \\
|sp^3\rangle &= |s\rangle\frac{1}{\sqrt{4}} + |p_z\rangle\sqrt{\frac{3}{4}} = |s\rangle\cos(60^\circ) + |p_z\rangle\sin(60^\circ) \\
|p_z\rangle &= |s\rangle\frac{1}{\infty} + |p_z\rangle\sqrt{\frac{\infty - 1}{\infty}} = |s\rangle\cos(90^\circ) + |p_z\rangle\sin(90^\circ)
\end{aligned}
$$

HYBRID ORBITALS BETWEEN S AND P STATES

Thus the building principle is

$$
|sp^n\rangle = |s\rangle\frac{1}{\sqrt{n+1}} + |p_z\rangle\sqrt{\frac{n}{n+1}} \tag{4.5}
$$

so that a pure s-orbital corresponds to a sp^0 hybrid and a pure p-orbital corresponds to a sp^∞ hybrid.

The building principle also allows us to point a hybrid orbital into an arbitrary direction, since we have already learned in Eq. 4.4 how to direct a p-orbital.

4.4 Hybrid orbitals and structure

Hybridization is an important factor determining the structure of an atom in a molecular or crystalline environment.

There are two main factors determining the structure around an atom.

1. Depending on the number of electrons, the formation of hybrid orbitals usually costs energy. This energy is called **promotion energy** and is shown in Tab. 4.1. The promotion energy depends on the number of electrons on the system.
2. The second factor is that an admixture of s-orbitals into the p-orbitals concentrates the orbital more on the bond, and thus leads to a stronger bond.
3. An atom can increase the number of bonds by introducing three-center bonds involving pure p-orbitals. This limits the maximum hybridization.

We will discuss these effects individually, before extracting general rules that allow us to estimate the local structure around an atom.

Promotion energy The promotion energy must be offset by bond formation. One can see that empty dangling bonds are preferably pure *p* orbitals, while filled dangling bonds, so-called lone pairs prefer to be in a hybrid orbital. This explains for example, why H_2O is not linear: it has two lone pairs on the oxygen atom, and therefore prefers sp^3 hybrids with a tetrahedral angle. In contrast, MgH_2, haveing two empty dangling bonds, prefers a linear configuration.

Bond-strength

Now we can determine the binding energy as function of the number of bonds. In order to arrive at a general statement, let us consider each bonding partner as a hydrogen atom. Instead of a hydrogen atom, we could also think of hybrid orbitals from potential bonding partners. Considering only a single type of bonding partners allows us to concentrate on the general characteristics.

The hopping integral of a central s-orbital with a hydrogen atom shall be $t_{ss\sigma}$ and that of a properly directed central p-orbital with a hydrogen shall be $-t_{sp\sigma}$. From these parameters we can derive the hopping integrals between the hybrid orbitals with the bonding partner using Eq. 4.5 for the sp-hybrid orbitals on the central atom. In order to compare the strength between different bond, we need an estimate for the relative strength of the hopping integrals $t_{ss\sigma} = 1.4\frac{\hbar^2}{md^2}$ and $t_{sp\sigma} = -1.84\frac{\hbar^2}{md^2}$, which we adopt from Eq. 3.13. Here d is the bond length.

$$
\begin{aligned}
\langle C_s|\hat{H}|H_s\rangle &= \langle C_s|\hat{H}|H_s\rangle = t_{ss\sigma} \approx 1.4\frac{\hbar^2}{md^2}\\
\langle C_{sp^1}|\hat{H}|H_s\rangle &= \frac{1}{\sqrt{2}}\langle C_s|\hat{H}|H_s\rangle + \frac{1}{\sqrt{2}}\langle C_p|\hat{H}|H_s\rangle = \frac{1}{\sqrt{2}}t_{ss\sigma} - \frac{1}{\sqrt{2}}t_{sp\sigma} \approx 2.29\frac{\hbar^2}{md^2}\\
\langle C_{sp^2}|\hat{H}|H_s\rangle &= \frac{1}{\sqrt{3}}\langle C_s|\hat{H}|H_s\rangle + \sqrt{\frac{2}{3}}\langle C_p|\hat{H}|H_s\rangle = \frac{1}{\sqrt{3}}t_{ss\sigma} - \sqrt{\frac{2}{3}}t_{sp\sigma} \approx 2.31\frac{\hbar^2}{md^2}\\
\langle C_{sp^3}|\hat{H}|H_s\rangle &= \frac{1}{\sqrt{4}}\langle C_s|\hat{H}|H_s\rangle + \sqrt{\frac{3}{4}}\langle C_p|\hat{H}|H_s\rangle = \frac{1}{\sqrt{4}}t_{ss\sigma} - \sqrt{\frac{3}{4}}t_{sp\sigma} \approx 2.29\frac{\hbar^2}{md^2}\\
\langle C_p|\hat{H}|H_s\rangle &= \langle C_p|\hat{H}|H_s\rangle = -t_{sp\sigma} \approx 1.84\frac{\hbar^2}{md^2}
\end{aligned}
$$

	0	1	2	3	4	5	6	7	8
s,p^3	0	0	0	0	0	0	0	0	0
sp,p^2	0	$\frac{1}{2}$	1	$\frac{1}{2}$	0	0	0	0	0
sp^2,p	0	$\frac{2}{3}$	$\frac{4}{3}$	1	$\frac{2}{3}$	$\frac{1}{3}$	0	0	0
sp^3	0	$\frac{3}{4}$	$\frac{3}{2}$	$\frac{5}{4}$	1	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	0

Table 4.1: Promotion energy for different sets of hybrid orbitals in units of $\epsilon_p - \epsilon_s$ for different numbers of electrons.

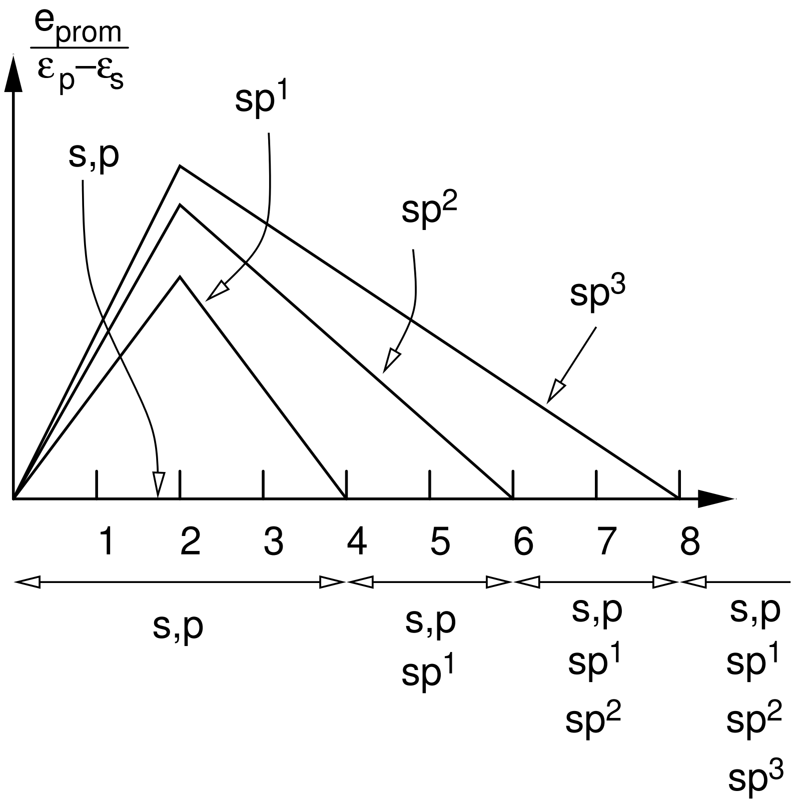


Fig. 4.6: Promotion energy for the various hybridizations as function of the number of electrons as derived in table 4.1. The number of electrons is obtained by counting all electron pairs involved with chemical bonds with the central to the central orbital. Below the horizontal axis the hybridizations with lowest energies are indicated.

It is apparent that the bond involving *sp*−hybrids are stronger than those with either s orbitals or p-orbitals. This is because the hybrids are concentrated in the bond direction, while s and p-orbitals point equally in both directions. Figure 4.7 shows the energy difference for bond formation relative to that of the unhybridized orbitals.

Three-center bonds

In addition to the simple two-center bonds, we consider also three-center bonds. At the central atom of a three center bond, there is usually a pure *s*−type orbital or a pure *p*−type orbital, because those can form equally strong bonds in two opposite directions.

In a three-center bond, the binding energy per partner is about 3/2 that of two individual bonds.[4] Thus adding another bonding partner to an existing bond to form a three-center bond is like forming a bond that is half as strong as a regular bond.

Rules for the coordination

This allows us to deduce an approximate rule for the coordination.

---

[4] In a two-center bond, the bonding orbital is stabilized by $t^2/|\epsilon_2-\epsilon_1|$, while for a three-center bond the stabilization is $\sqrt{2}t^2/|\epsilon_2-\epsilon_1|$. Thus we can attribute a bond-strength of $(\sqrt{(2)}-1)t^2/|\epsilon_2-\epsilon_1|$ to the second bond. Thus this additional bond is about one-half as strong as the first bond.

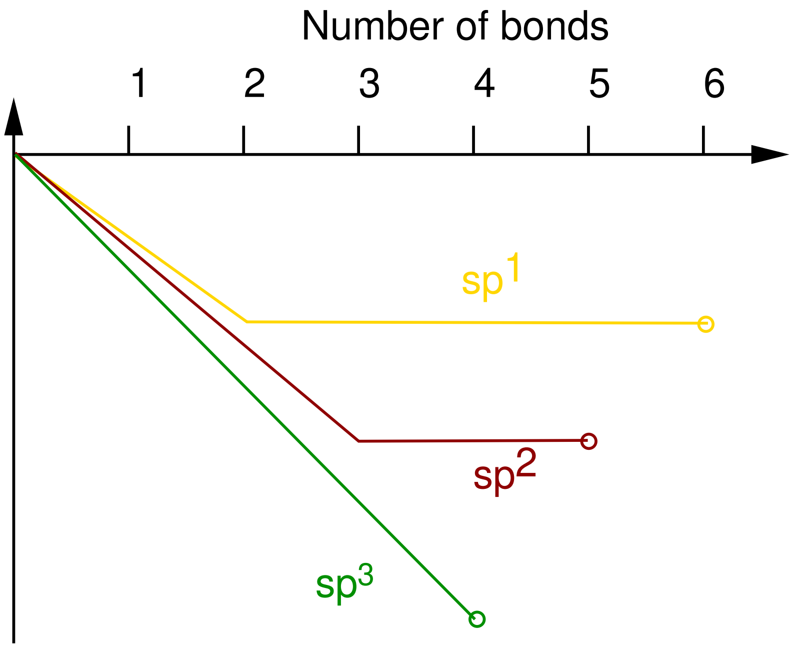


Fig. 4.7: Bond energy relative to the unhybridized situation for different hybridizations as function of the number of bonds.

RULES FOR SP-BONDED ATOMS

- We attribute the electron pairs from the bonds to the central atom. The number of electrons excludes certain hybridizations as shown in Fig. 4.6
- From the set of allowed hybridizations we determine the largest bond energy compatible with the number of bonds from diagram Fig. 4.7

4.4.1 Hydrides of main group atoms

We will apply these rules to the hydrides of the main group elements. The s-orbital of a hydrogen atom is a representatives of a general dangling bond of a neighboring atom.

Let us consider the case with eight electrons, for which the promotion energy does not exclude any hybridization. Because the bonds with sp-hybrids are strongest, we form first sp^3-hybrids[5] This leads to the series HF, H_2O (water), NH_3(ammonia) and CH_4(methane). As shown in table 4.2, all these molecule have bond-angles close to the tetrahedral bond angle, namely 109.5°. For the heavier elements though, the angle becomes smaller and approaches 90°. This effect can be attributed to an admixture of unoccupied d-orbitals as it is demonstrated in fig. 4.8. The effect clearly sets in with the 3rd period, because this is the first valence shell containing d-electrons with the same principal quantum number.

In order to see how the rule works one may consider the systems with the same number of bonds but different electron number, namely BH_3 and NH_4. BH_3 has 6 electrons. Hence the argument of minimizing the promotion energy excludes the sp^3 hybridization. Since one of the orbitals remains empty it is best to have one orbital lying as high as possible. Note, that the sum of the orbitals

---

[5] According to our estimate the sp^2 orbitals would be more favorable. However, the real structures indicate that sp^3 orbitals are stronger.

HF	116°	H_2O	104.45°	NH_3	107.8°	CH_4	109.5°
		H_2S	92.1°	PH_3	93.5°	SiH_4	109.5°
		H_2Se	91°	AsH_3	91.8°	GeH_4	109.5°
		H_2Te	90°	SbH_3	91.7°	SnH_4	109.5°

Table 4.2: Bond angles of the 8-electron hydrides. In the first row, the bond angles are close to the tetrahedral bond angle of 109.5°, consistent with sp^3 hybridization. For the heavier elements the bond angles approach 90°. One possible explanation for the 90° bond angle is an admixture of the unoccupied d-electron shell as illustrated in fig. 4.8.

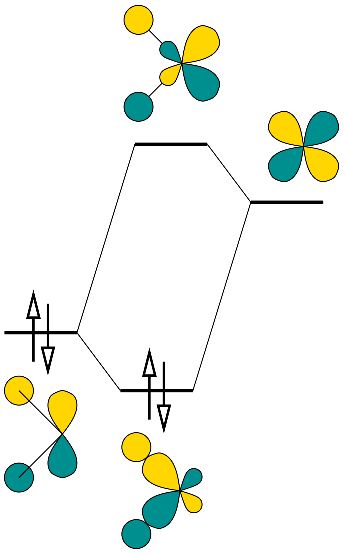


Fig. 4.8: Demonstration of the admixture of empty d-orbitals to the filled bonding orbitals. The result is a reduction of the bonding angle towards 90°. This effect is only present in the 3rd and higher periods because the 2nd shell does not have d-orbitals.

energies is independent of the hybridization. Next we select that hybridization with the largest bond energy from figure 4.7. Thus we arrive at the sp^2 hybridization, which indicates that the molecule is planar and has bond angles of 120°. Because the electron count and the number of bonds are the same, this conclusion also holds for the so-called carbocation CH_3^+, and important intermediate in organic reactions.

If we add two electrons, the situation changes, and we obtain an umbrella like structure consistent with sp^3 hybridization.

One special coordination is the 5-fold coordination. Atoms that form bonds to more partners than expected from their valency are called hypervalent atoms. Here, a trigonal bi-pyramid is formed, with three equatorial bonds with bond angles of 120° on the one hand and one three center bond between the two axial hydrogen atoms with the p_z orbital.

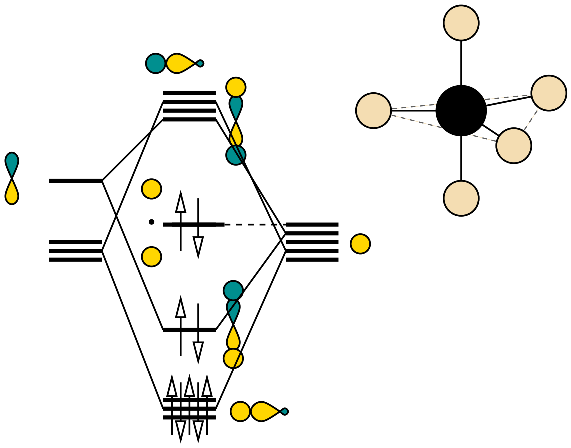
electrons	bonds	binding
0	0	s
2	0	s, like helium
2	1	$ss\sigma$ bond
2	2	s-type three-center bond
4	0	s
4	1	sp^1 bond and sp^1-type lone pair as back bond
4	2	two sp^1 bonds, linear molecule
6	0	sp^1
6	1-3	sp^2 bonds, bond angle 120°
8	0	sp^2, like noble gases
8	1-4	sp^3, tetrahedral bond angle
8	5	three equatorial sp^2 bonds and one axial three-center bond
8	6	three orthogonal three-center bonds, octahedron
10	5	three equatorial sp^2 bonds and one axial three-center bond
10	6	three orthogonal three-center bonds and one filled non-bond
12	6	three orthogonal three-center bonds and two filled non-bonds
14	6	three orthogonal three-center bonds and three filled non-bonds

Table 4.3: Preferred hybridization of an anion as function of number of electons and bonds. The electron pair of a bond is added to the electron count. For the cases with zero bonds, we consider the configuration favorable for a slightly positive charge, given that the attempt to form a bond with an anion is likely to remove electrons. Filled non-bonds refer to the non-bonding orbital of a three-center bond.

compound	N_e	Hybrid	coordination
LiH	2	s,p	linear
BeH_2	4	sp^1	linear
BH_3	6	sp^2	trigonal
CH_3^+	6	sp^2	trigonal
CH_4	8	sp^3	tetrahedral
NH_4^+	8	sp^3	tetrahedral
NH_3	8	sp^3	umbrella
H_3O^+	8	sp^3	umbrella
H_2O	8	sp^3	bent
HF	8	sp^3	linear
CH_5^-	10	sp^2	trigonal bipyramidal
PH_5	10	sp^2	trigonal bipyramidal

Note that the hybrid orbitals allow us to estimate the atomic structure. However, they do not tell us about the atomic orbitals. In order to obtain the eigenstates, we need to form symmetrized orbitals from the bonding and anti-bonding orbitals with the hybrid orbitals. The transformation into eigenstates, however, does not change the energy of the system, because only occupied orbitals are transformed into each other and unoccupied orbitals mix with each other.

The notation of hybrid orbitals thus shortcuts the complete diagonalization of the Hamiltonian, which is sufficient to estimate total energies and structures.



4.5 Coordination of transition metal compounds

While main group elements are the work horse of chemistry and material science, the salt in the soup are the transition metals. They are very flexible in the way, they form bonds. Therefore they play an important role in catalysts. They are also involved in the most interesting class of solid materials namely transition-metal oxides. These can be metals and insulators, ferromagnets or antiferromagnets, piezoelectrics, ferroelectrics, and even super superconductors.

There are a number of coordinations that play a special role for transition metal compounds.

- tetrahedral (4 neighbors)
- trigonal bipyramid (5 neighbors)
- octahedral (6 neighbors)
- square planar (4 neighbors)

When trying to understand bonding in transition metal compounds one should always concentrate on the anti-bonding states.

Because transition metals are electropositive, the d-orbitals normally lie above the ligand orbitals. Hence the bonding states are mostly concentrated on the ligands, while the anti-bonding orbitals have mostly d-character. because there is a bond for every anti-bond, we can estimate structural stability already from the anti-bonds.

The *d*-orbitals are usually expressed by cubic or real harmonics. There are five d-orbitals named $d_{3z^2-r^2}$ $d_{x^2-y^2}$, d_{xy}, d_{xz}, d_{yz}. The index gives the shape of the orbitals and neglects only the prefactors.

4.5.1 Square-planar coordination

In the square planar coordination the four ligands can form directed bonds to a single orbital, namely $d_{x^2-y^2}$. Square-planar complexes are typical for a d^8 configuration, that is there are 8 d-electrons in the d-shell.

4.5.2 Tetrahedral coordination

The tetrahedral coordination is most favorable for d^4 complexes.

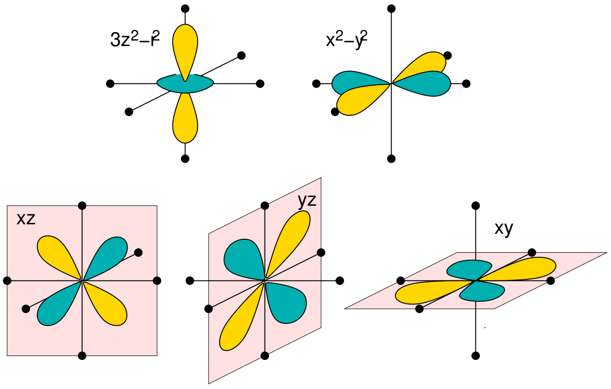


Fig. 4.9:

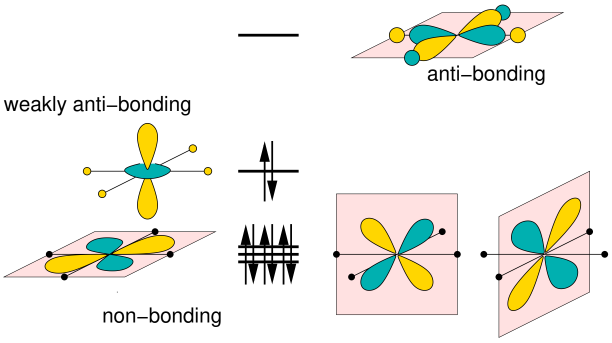


Fig. 4.10: Orbital scheme for square-planar transition metal complexes.

The d-levels split into a low-lying pair of non-bonding orbitals and a triple of antibonding orbitals. The pair is denoted by e_g and the triple is denoted by t_{2g}.

The letter e in e_g stands for “entartet”, the german word for degenerate, and the subscript g gerade, the german word for “even”. What is meant is that the orbitals are symmetric unter inversion. s- and d-orbitals are symmetric under inversion while p and f orbitals are antisymmetric. For the latter the g would be replaced by “u”, that is “ungerade” the german word for “odd”.

The letter t in t_{2g} stands for “triple”, that is three-fold degenerate. The subscript g stands again for “gerade”. I have not found a meaning for the subscript “2”.

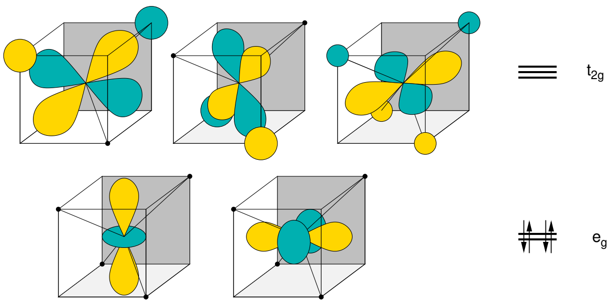


Fig. 4.11: Orbital scheme for tetrahedral transition metal complexes.

4.5.3 Octahedral coordination

The octahedral coordination is most favorable for d^6 coordination. It is probably the most common coordination for late transition metal compounds.

Characteristic for the octahedral environment of a transition metal is that there are two unoccupied anti-bonding levels e_g and three occupied non-bonding levels t_{2g}.

Thus the order of e_g and t_{2g} orbitals is opposite to that in the tetrahedral coordination.

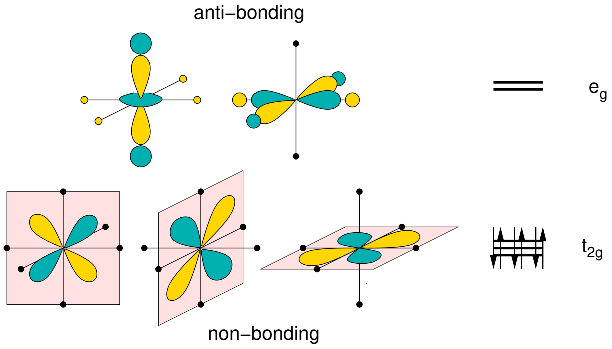


Fig. 4.12: Orbital scheme for octahedral transition metal complexes.

An example for octahedral bonding are the perovskites.

4.5.4 s-p-d hybrids

It is also possible to form hybrid orbitals between s-, p-, and d-orbitals. I will not discuss those for the following reason. Orbitals with the same principal quantum number have a range. More precisely, the radial parts of the corresponding orbitals in the bonding region is similar. Therefore, if we form hybrid orbitals between orbitals with the same principal quantum number, the shape of the hybrid orbital dominated by the angular part, that is the spherical harmonics.

The valence *d* orbitals, however, have a different principal quantum number than the valence s- and p-orbitals. The shape of the radial parts is therefore very different and the shape of the hybrid orbital depends also on the radial parts.

The same is true for *f* orbitals, which have an extent that is similar to core states, even though their energy is in the valence region.

This is the reason for not considering hybrid orbitals build from states with different principal quantum number.

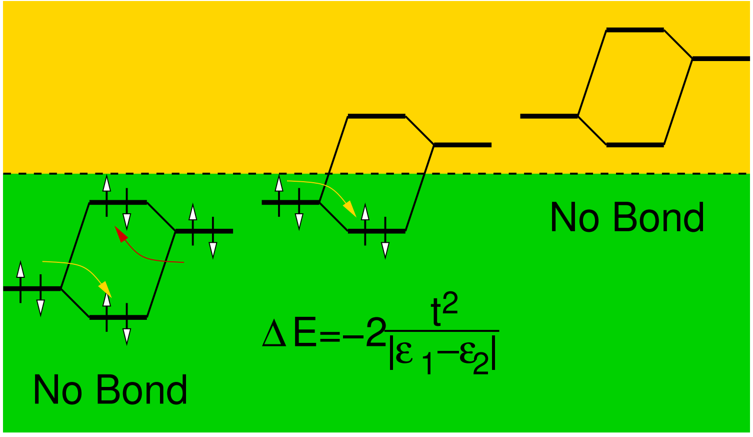
Chapter 5

Frontier orbitals

Often we are interested in reactivity of a molecule or a material. We would like to find out those parts of the molecule that are most reactive, and we would like what kinds of reactions these "hot spots" can undergo. **Frontier orbital** theory is a simple tool that allows us to make exactly such an analysis. Frontier orbital theory was invented by Kenichi Fukui who received in 1981 the Nobel price in chemistry for this contribution[5].

The frontier orbital concept, in its most simple form, says that the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO) are usually those that are responsible for the reactivity.

The underlying idea is that only the hybridization of filled with empty orbitals leads to a stabilization of the system. Because the interaction is largest if the energetic separation between the participating orbitals is smallest, it is obvious that the orbitals near the Fermi level, contribute mostly to the interaction.



One differentiates nucleophilic and electrophilic sites at a molecule.

- an **electrophile** is characterized by a low-lying empty orbital, which can undergo an interaction with an occupied orbital. "Electrophile" means electron-loving.
- a **nucleophile** is the opposite of an electrophile, and is characterized by a high-lying filled orbital. This orbital can undergo an interaction with an electrophile. "Nucleophile" means nucleus-loving. Consider for example a proton, which is a nucleus. A proton has an empty orbital, and thus interacts best with a filled orbital. Thus a filled orbital is nucleus-loving.

There are also sites that can be both, electrophilic and nucleophilic at the same time, if it has both a low-lying empty and a high-lying filled orbital.

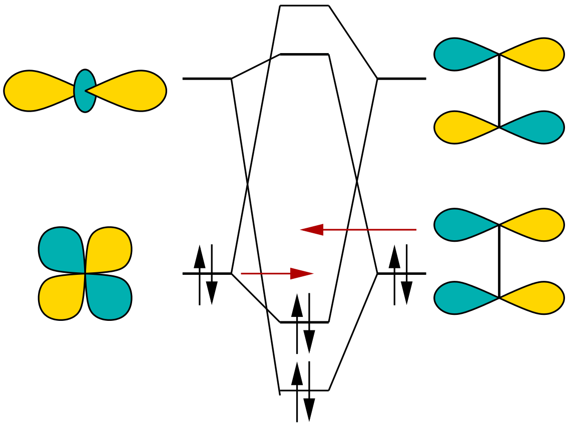


Fig. 5.1: The concept of donation and back donation, a frequent binding scheme between transition metals and double bonds.

It should be noted that it is not always the HOMO and the LUMO. One should always consider all states close to the HOMO and the LUMO, which are accessible from the outside.

5.1 Determine strong bonds

In the first step we determine those parts of the molecule which is not reactive, which is the σ-bond network. Each σ-bond receives two electrons, which are subtracted from the total number of valence electrons.

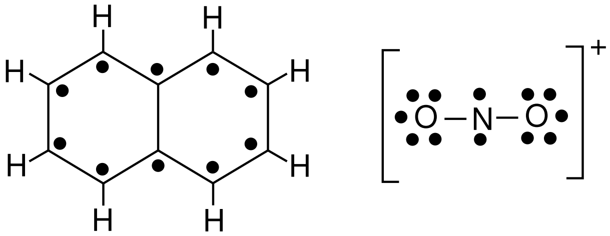


Fig. 5.2: Naphthalene (left) and Nitronium (right) with σ-bonds and remaining electrons.

5.1.1 Nitronium ion

Let us first concentrate the nitronium ion or nitryl ion NO_2^+. We have drawn the molecule as bent, but currently this is merely a guess. The total number of electrons is 16. The sigma bonds consume 4 electrons. Thus, we have to distribute the remaining 12 electrons onto the π orbitals.

Since the molecule is planar there is certainly a mirror plane, which separates the p orbitals perpendicular to the plane. The π system forms a three-center bond and thus has one bonding one anti-bonding state and a non-bonding state in the middle. The bonding orbital can consume two orbitals and the non-bonding can consume up to four orbitals.

In the orbital system that is symmetric with respect to the mirror plane, there would be one sp^2 orbital on nitrogen and two orbitals each on oxygen, which can be either two sp^2 hybrids, or a sp^1

hybrid and a *p*-orbital in the plane. If we form lone pairs out of all these orbitals, that is we fill them with electrons, we can deposit 10 electrons into the lone pairs.

Since the bonding state of the pi-system is clearly more stable than the lone-pairs, it will be filled first. Since the remaining orbitals are sp^2 hybrids, they are lower in energy than the non-bonding orbital, which has the energy of a p-electron. Thus we would judge that there are ten electrons in sp^2 lone pairs and two in the three center bond of the π system

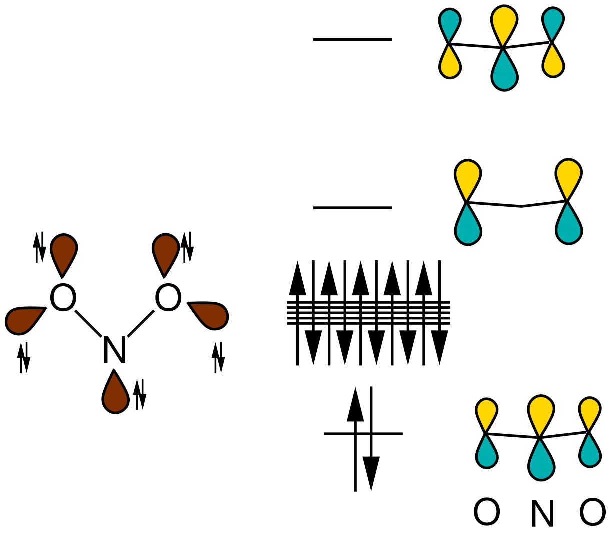


Fig. 5.3: Hybrid orbitals and energy level scheme of Nitronium assuming a bent configuration and non-interacting sp^2 hybridization.

This looks ok, but some aspects are incorrect. If the π-triple bond would have the same weight on all three atoms, we would obtain a charge of $5\frac{2}{3}$ on the oxygen atoms and a charge of $4\frac{2}{3}$ on the nitrogen atom.The formal charges would be $N^{+\frac{1}{3}}O_2^{+\frac{1}{3}}$. Worse, the bonding state of the triple bond is localized mostly on the nitrogen, which would make the nitrogen more negative than the oxygen atoms. This however would contradict the trend of electronegativities. We have changed the sp^2 hybridization to a sp^1 hybridization. There is a small cost, namely $\frac{1}{2}(\epsilon_p - \epsilon_s)$ in promotion energy.

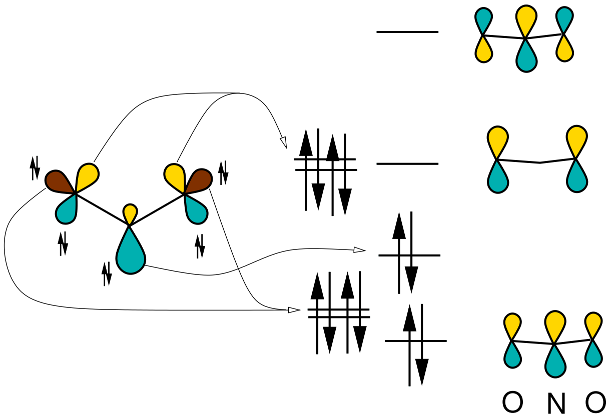


Fig. 5.4: Hybrid orbitals and energy level scheme of Nitronium assuming a bent configuration and sp/p configuration on the terminal oxygen atoms. Observe the interaction of the p-orbitals of the oxygen atoms with the back lobes of the sp^2 hybrid on nitrogen.

However when we consider the interaction including the back lobe of the sp^2 orbital on nitrogen, we see that we can form an additional bond. The system of the three orbitals behaves again like a three center bond, which has a non-bonding and a bonding orbital. The bonding orbital is strengthened, if the molecule becomes more planar, because this increases p-character of the orbital on nitrogen and thus strengthens the π bond character. Thus we can see that nitrogen dioxide is a linear molecule.

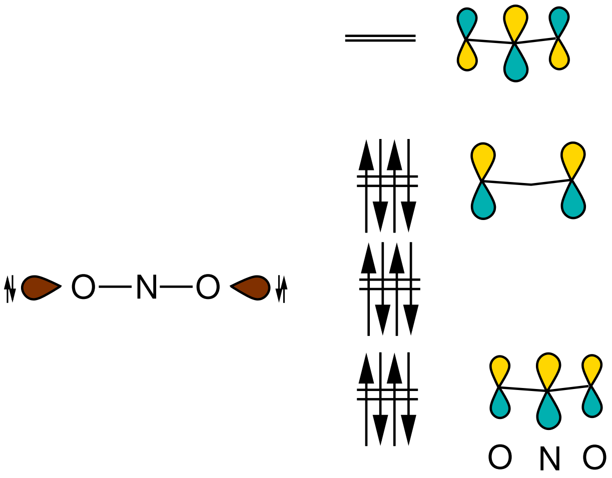


Fig. 5.5: Hybrid orbitals and energy level scheme of Nitronium assuming a linear configuration.

The nitronium ion is isoelectronic with carbon dioxide, which also is a linear molecule. As we add electrons, we expect the molecule to become bent, because we fill the anti-bonding orbital of the three center bond. Experimentally, nitrogen dioxide NO_2^0 is bent with an angle of 134° and is a radical, that is it has one unpaired electron. The nitrite ion NO_2^- has an angle of 115°, close to the 120° of an sp^2 hybrids on the nitrogen atom.

5.1.2 Naphthalene

In order to understand the reactivity of naphthalene we need to find out the nature of the HOMO and the LUMO.

Drawing the sigma-bond network is simple for naphthalene. The only complicated point is to determine the π-states. We exploit the two mirror planes of the molecule to break down the orbitals into symmetry groups, which is shown on the left side of Fig. **??**

We obtain three groups with three orbitals and two groups with 2 orbitals.

- for the sets with three orbitals we observe that the middle one can interact with the left and the right orbital, but that there is no interaction between the left and the right orbital. Thus, these three orbitals behave like a three center bond. so that it is easy to draw the bonding, the non-bonding and the anti-bonding orbital.
- for the sets with two orbitals, we obtain a bonding and an anti-bonding state.

In order to estimate the energetic order, we simply add the number of bonding and non-bonding interactions and divide by the number of p orbitals in the wave function. Note that this approximation overestimates the stability of the three center bond, but for our qualitative argument this estimate is sufficient.

We detect immediately the HOMO and the LUMO of the molecule, which have energy estimates of $\pm\frac{1}{4}$ in units of t.

After we determined the relevant orbitals, let us take a closer look at HOMO and LUMO. Our simplistic estimate ignored the fine details. Let us concentrate on the HOMO: The bonding pair of

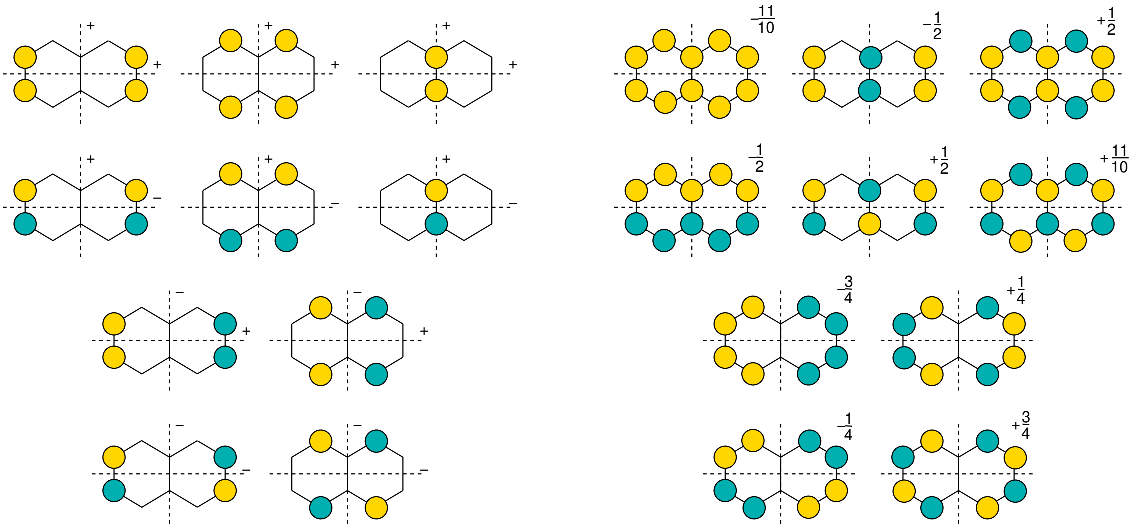


Fig. 5.6: Left side: symmetrized basis functions for the π orbital system of naphthalene. Right side: approximate eigenstates for the π-electron system naphthalene.

π orbitals is not in an symmetric environment. It has an anti-bond on the one side, and no orbital on the other. The orbital can lower its energy by weakening the anti-bond. Thus it places more weight onto the orbital next to the missing orbital.This will be crucial for the selectivity of the reaction

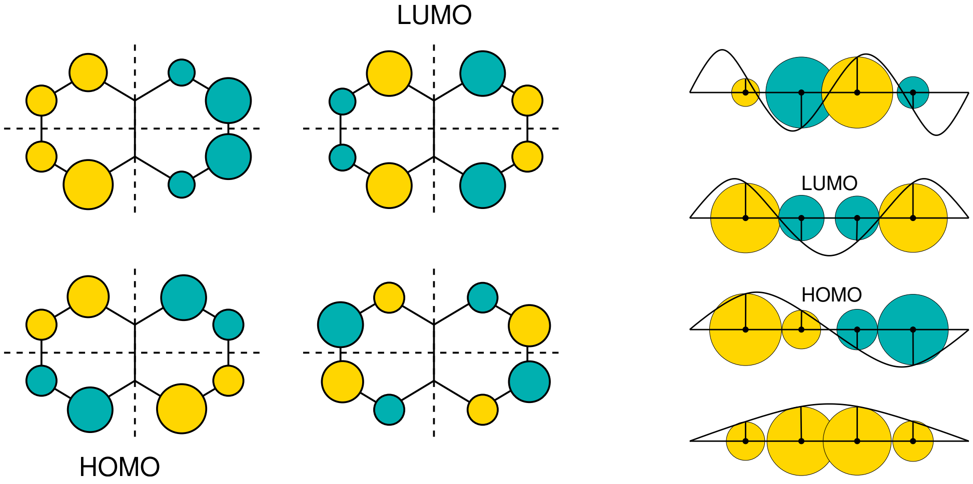


Fig. 5.7: The HOMO and the LUMO of Naphthalene along with the respective partner wave functions of the same symmetry group. The figure on the right shows how one can estimate the amplitude of the orbitals from sinus functions.

5.1.3 Nitridation of Naphtalene

As a positive ion, the nitronium ion certainly wants to gain more electrons. Thus we look for its LUMO, which is one of the triple-anti-bonds. This orbital has the largest weight on the nitrogen atom. The LUMO of the nitronium ion wants to interact with the HOMO of naphthalene, which has its largest weight next on the carbon atom next to the central bond. Thus we predict the nitronium ion to attack preferentially this carbon atom.

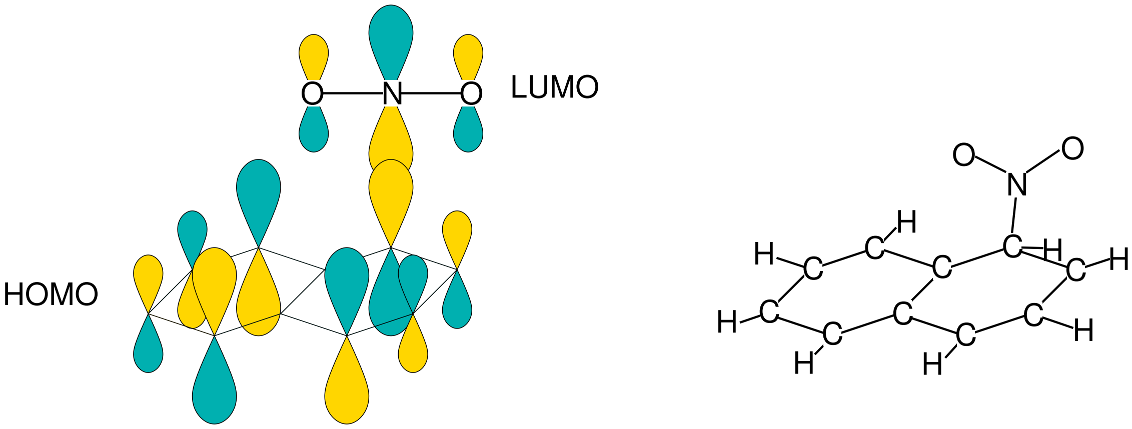

O
N
O
LUMO
HOMO
O
N
O
H
C
H
H
C
H
H
C
C
H
C
C
C
C
H
H
C
H

Fig. 5.8: The electrophilic addition of nitronium to naphthalene. and final structure.

Chapter 6

From bonds to bands

So far we have employed a local picture of bonding. In solid state physics another viewpoint became popular, based on the concept of reciprocal space and band structures.

In order to approach the concepts in a simple way we again choose a simple model system, namely the jellium model.

6.1 The Jellium model

For the electron gas the model system is the **free electron gas** or the **jellium model**. The jellium model consists of electrons and a spatially constant, positive charge background, which ensures that the overall system is neutral.

Because of translational symmetry, the potential is spatially constant. Thus the one-particle states are simply plane waves.

$$\phi_{\vec{k},\sigma}(\vec{r},\sigma') = \frac{1}{\sqrt{\Omega}} \mathrm{e}^{i\vec{k}\vec{r}} \delta_{\sigma,\sigma'} \tag{6.1}$$

We consider here states in a very large, but finite, volume Ω such as the universe. The states are normalized to one within this volume.

This is an unusual notation for an ideally infinite system. Using this notation we avoid having to distinguish discrete and continuous sections of the spectrum and different normalization conditions for localized and extended states. A disadvantage is that extended states have a nearly vanishing amplitude, and that the spacing in k-space is extremely small. For a finite volume, only discrete values for $\vec{k}$ are allowed. We will later see that in the final expression the factors Ω from the normalization can be translated into a volume element in k-space, so that the sums can be converted into integrals.

The dispersion relation of free electrons forms a parabola

$$\epsilon_{\vec{k},\sigma} = V_0 + \frac{(\hbar k)^2}{2m_e}$$

where V_0 is the value of the potential.

The jellium model approximates real metals surprisingly well, even though the potential of the nuclei is far from being a constant. In a hand-waving manner we can say that the valence electrons are expelled from the nuclear region by the Pauli repulsion of the core electrons, so that they move around in a region with fairly constant potential. The true story is a little more subtle though... In Fig. 6.1 the band structure of free and non-interacting electron gas is shown in a unit cell corresponding to alumina and compared to the calculated band structure of alumina. We can see that the free electron gas already provides a fairly good description of some realistic systems.

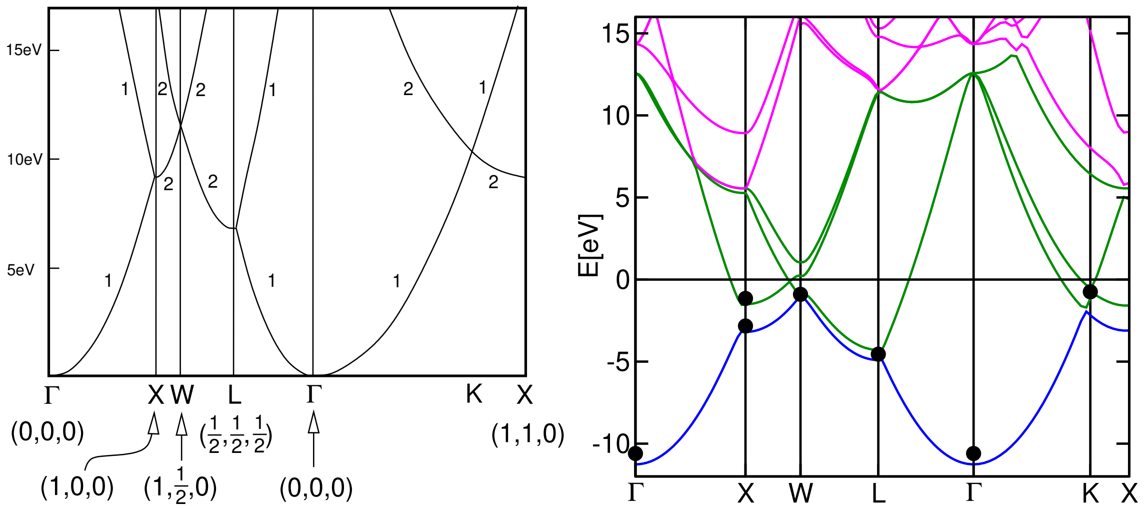


Fig. 6.1: Band structure of free, non-interacting electrons. The lattice is an fcc-cell with a lattice constant of 4.05 Å corresponding to aluminum. The high symmetry points are given in units of $\frac{2\pi}{a_{lat}}$. The numbers indicate the degeneracy beyond spin-degeneracy. On the right-hand side, the band structure of aluminum is shown in comparison.

6.2 Density of States

6.2.1 Motivation

While the dispersion relation contains a wealth of information, that is very important for transport problems or the interaction of quasi-particles, for many purposes the k-dispersion is not relevant. I those cases the additional information of the k-dispersion obscures the truely important information.

Thermodynamic potentials such as the free energy can be written, for non-interacting, identical particles as a sum over energy levels.

$$\langle A\rangle_{\beta,\mu}=\sum_n f_{T,\mu}(\epsilon_n)\langle\psi_n|\hat{A}|\psi_n\rangle=\int d\epsilon\; f_{T,\mu}(\epsilon)\underbrace{\sum_n\delta(\epsilon-\epsilon_n)\langle\psi_n|\hat{A}|\psi_n\rangle}_{D_A(\epsilon)}$$

where

$$f_{T,\mu}(\epsilon)=\left(1+\mathrm{e}^{\frac{1}{k_BT}(\epsilon_n-\mu)}\right)^{-1}$$

is the Fermi distribution function.

The density of states separates "thermodynamic information" such as temperature and chemical potential from the "system intformation" which refers to the energies and matrix elements of individal states.

One can go one step further and separate out the observable A, if one introduces a local basisset $|\chi_\alpha\rangle$ and corresponding projector functions $\langle\pi_\alpha|$, that obey the bi-orthogonality condition $\langle\pi_\alpha|\chi_\beta\rangle=\delta_{\alpha,\beta}$. For an orthonormal basisset one may choose the projector functions equal to the basis-functions themselves. For non-orthonormal basissets, the projector functions carry another inverse overlap matrix. One can verify that

$$|\psi\rangle=\sum_\alpha|\chi_\alpha\rangle\langle\pi_\alpha|\psi\rangle\qquad\text{if }\langle\pi_\alpha|\chi_\beta\rangle=\delta_{\alpha,\beta}$$

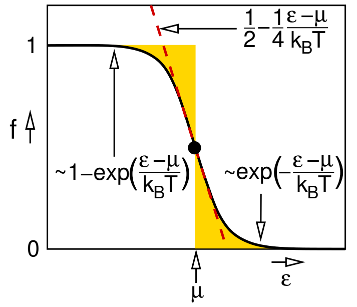


Fig. 6.2: Fermi distribution $f(\epsilon)=(1+\mathrm{e}^{\frac{1}{k_BT}(\epsilon-\mu)})^{-1}$.

Insertion into the expression for the expectation value above yields

$$\begin{aligned}\langle A\rangle_{\beta,\mu} &= \int d\epsilon\; f_{T,\mu}(\epsilon)\sum_n \delta(\epsilon-\epsilon_n)\sum_{\alpha,\beta}\langle\psi_n|\pi_\alpha\rangle\langle\chi_\alpha|\hat{A}|\chi_\beta\rangle\langle\pi_\beta|\psi_n\rangle\\ &= \int d\epsilon\; \sum_{\alpha,\beta} f_{T,\mu}(\epsilon)\underbrace{\sum_n\langle\pi_\beta|\psi_n\rangle\delta(\epsilon-\epsilon_n)\langle\psi_n|\pi_\alpha\rangle}_{D_{\beta,\alpha}(\epsilon)}\langle\chi_\alpha|\hat{A}|\chi_\beta\rangle\end{aligned}$$

In this way we have managed a division of the expression into

- thermodynamic information, i.e the Fermi distribution function
- system-dependent information, the density of states

$$D_{\alpha,\beta}(\epsilon)\overset{\text{def}}{=}\sum_n\langle\pi_\beta|\psi_n\rangle\delta(\epsilon-\epsilon_n)\langle\psi_n|\pi_\alpha\rangle \tag{6.2}$$

  and the

- observable information, namely the matrix element $\langle\chi_\alpha|\hat{A}|\chi_\beta\rangle$.

Most important is the **total density of states**, where the operator A is the unit operator.

$$D(\epsilon)=\sum_n\delta(\epsilon-\epsilon_n)$$

Often also **projected density of states** are used, where the density of states is projected onto a certain orbital

$$D_{\alpha,\alpha}(\epsilon)=\sum_n\delta(\epsilon-\epsilon_n)\langle\psi_n|\chi_\alpha\rangle\langle\chi_\alpha|\psi_n\rangle$$

In order to investigate bonding one also uses off-diagonal elements of the density of states matrix. There are several variants: Mulliken's overlap populations and the Crystal orbital overlap contributions (COOP) of Hoffman and Hughbanks use the overlap matrix elements A variant more directly related to the energetics are the and the **Crystal Orbital Hamilton Populations (COHP)**[13].

$$COHP_{\alpha,\beta}(\epsilon)=D_{\alpha,\beta}(\epsilon)\langle\chi_\beta|\hat{H}|\chi_\alpha\rangle$$

The energetic information contained in the density of states is important because it provides some insight into how robust a certain expectation value is against perturbations of the system.

6.2.2 Density of States for extended systems

For molecules, the total density of states is simply a series of Delta functions. For extended systems this is no more true, but the density of states becomes a continuous function.

The one-particle density of states can be obtained from the dispersion relation $\epsilon(p)$. Independent particles in a constant potential have a conserved momentum $\vec{p} = \hbar\vec{k}$. This implies that they can be described by plane waves

$$\Psi(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k}\vec{r}}$$

Note that even electrons in a periodic potential, that is electrons ins a crystal, have a conserved momentum. The wave function is not a plane wave, but a plane wave modulated by some periodic, p dependent function, as seen in the **Bloch theorem**. These dispersion relations are called **band structure** and can be calculated with first-principles calculations. Also lattice vibrations in a crystal can be classified by their wave vector, resulting in a phonon band structure.

If we consider a system with periodic boundary conditions in a a box with side-lengths L_x, L_y, L_z, the states are quantized, so that $k_i L_i = 2\pi n_i$ where n_i is an arbitrary integer. Thus there are only states with $\vec{k} = (\frac{2\pi}{L_x} i, \frac{2\pi}{L_y} j, \frac{2\pi}{L_z} k)$, where i, j, k are arbitrary integers. The volume V of the box is $V = L_x L_y L_z$.

Thus the density of states is

$$D(\epsilon) = \sum_{i,j,k} \delta\Big(\epsilon(\underbrace{\frac{2\pi\hbar}{L_x} i}_{p_x}, \underbrace{\frac{2\pi\hbar}{L_x} j}_{p_y}, \underbrace{\frac{2\pi\hbar}{L_x} k}_{p_z}) - \epsilon\Big)$$

We can attribute to each state a volume in k-space, namely

$$\Delta V_k = \frac{2\pi}{L_x}\frac{2\pi}{L_y}\frac{2\pi}{L_z} = \frac{(2\pi)^3}{V} \tag{6.3}$$

Using the relation $\vec{p} = \hbar\vec{k}$ we can convert this volume into a volume element in momentum space, namely

$$\Delta V_p = \frac{(2\pi\hbar)^3}{V} \tag{6.4}$$

If the size of the box, that is L_x, L_y, L_z, is made very large, the volume element attributed to a single state in momentum space becomes very small. Thus we can replace the sum by an integral, where ΔV_p is represented by d^3p.

$$\begin{aligned} D(\epsilon) &= \frac{V}{(2\pi\hbar)^3} \sum_{i,j,k} \underbrace{\frac{(2\pi\hbar)^3}{V}}_{\to d^3p} \delta\Big(\epsilon(\vec{p}_{i,j,k}) - \epsilon\Big) \\ &\overset{L_i \to \infty}{=} \frac{V}{(2\pi\hbar)^3} \int d^3p \, \delta\Big(\epsilon(\vec{p}) - \epsilon\Big) \end{aligned} \tag{6.5}$$

It is intuitively clear that the expression for the density of states is related to a surface integral over a surface of constant energy. This will be shown in the following.

In order to transform the expression for the density of states into a surface integral in momentum space, it is convenient to introduce the **number of states** $N(\epsilon)$.

The number of states is defined as the number of states that lie below the energy epsilon.

$$N(\epsilon) = \sum_{i,j,k} \theta(\epsilon - \epsilon(p_{i,j,k})) \overset{V \to \infty}{=} \frac{V}{(2\pi\hbar)^3} \int d^3p \, \theta(\epsilon - \epsilon(\vec{p})) \tag{6.6}$$

where $\theta(x)$ is the Heaviside function $\theta(x)$, which vanishes for $x < 0$ and is equal to unity for $x > 0$. The Heaviside function is related to the δ-function via

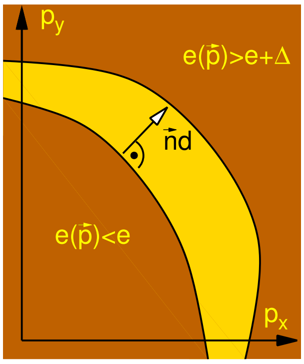
$$\theta(x) = \int_{-\infty}^{x} dx'\ \delta(x') \qquad \Leftrightarrow \qquad \partial_x \theta(x) = \delta(x) \tag{6.7}$$

This allows us to relate the number of states to the density of states Thus we obtain

$$\begin{aligned}\partial_\epsilon N(\epsilon) &\overset{\text{Eq. 6.6}}{=} \frac{V}{(2\pi\hbar)^3} \int d^3p\ \partial_\epsilon \theta(\epsilon - \epsilon(\vec{p})) \overset{\text{Eq. 6.7}}{=} \frac{V}{(2\pi\hbar)^3} \int d^3p\ \delta(\epsilon - \epsilon(\vec{p})) \\ &\overset{\text{Eq. 6.5}}{=} D(\epsilon)\end{aligned} \tag{6.8}$$

Here we show how the volume integral with the delta function can be converted into a surface integral.

$$D(\epsilon) \overset{Eq. 6.8}{=} \partial_\epsilon N(\epsilon) = \lim_{\Delta\to 0} \frac{V}{(2\pi\hbar)^3} \int d^3p\ \underbrace{\frac{\theta(\epsilon + \Delta - \epsilon(p)) - \theta(\epsilon - \epsilon(p))}{\Delta}}_{\to\delta(\epsilon-\epsilon(p))}$$



The integral over the difference of the two Heaviside function corresponds to the volume of a volume sheet, which is enclosed by the surfaces defined by $\epsilon(\vec{p}) = \epsilon$ and $\epsilon(\vec{p}) = \epsilon + \Delta$.

Let us calculate the distance of two points on the two surfaces. Let us pick one point $\vec{p}_0$ on the surface defined by $\epsilon(\vec{p}) = \epsilon$. The closest neighbor on the other surface $\vec{p}_1$ lies opposite to $\vec{p}_0$, that is $\vec{p}_1 = \vec{p}_0 + \vec{n} \cdot d$, where d is the distance of the points and $\vec{n} = \frac{1}{|\vec{\nabla}_p \epsilon|} \vec{\nabla}_p \epsilon$ is the normal vector of the surface. $\vec{p}_1$ lies on the other surface and therefore fulfills

$$\begin{aligned}\epsilon(\vec{p}_1) &= \underbrace{\epsilon}_{\epsilon(\vec{p}_0)} + \Delta \\ \Rightarrow \epsilon(\vec{p}_0) + \Delta &= \epsilon(\vec{p}_0 + \vec{n} \cdot d) \overset{Taylor}{=} \epsilon(\vec{p}_0) + \vec{n} \cdot d\ \vec{\nabla}_p \epsilon + O(d^2) \\ \Rightarrow d &= \frac{\Delta}{\vec{n}\vec{\nabla}_p \epsilon} = \frac{\Delta}{\frac{\vec{\nabla}_p \epsilon}{|\vec{\nabla}_p \epsilon|} \vec{\nabla}_p \epsilon} = \frac{\Delta}{|\vec{\nabla}_p \epsilon|}\end{aligned}$$

Thus we obtain the thickness d of the sheet. Note that the thickness depends on the position on the sheet. The volume element of the sheet can then be written as $d^3p = dA \cdot d$

Thus we can write the volume as

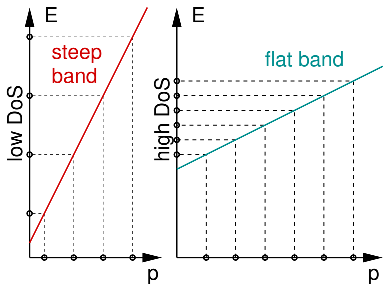
$$D(\epsilon) = \frac{V}{(2\pi\hbar)^3} \oint d^2A_p\ \frac{1}{|\nabla_p \epsilon|}$$

Thus we obtain the density of states from the dispersion relation as

ONE-PARTICLE DENSITY OF STATES PER VOLUME

$$g(\epsilon)\stackrel{\text{def}}{=}\frac{1}{V}D(\epsilon)=\int\frac{d^3p}{(2\pi\hbar)^3}\delta\left(\epsilon-\epsilon(\vec{p})\right)=\frac{1}{(2\pi\hbar^3)}\oint d^2A_p\,\frac{1}{|\vec{\nabla}_p\epsilon(p)|}\tag{6.9}$$

Note that $\vec{\nabla}_p\epsilon(p)$ is the velocity (group-velocity) of the particle. This follows from Hamilton's equation $\dot{x}_i=\frac{\partial H(\vec{p},\vec{x})}{\partial p_i}$. It also follows from the expression of the group velocity of a wave packet $v_i=\frac{\partial\omega(\vec{k})}{\partial k_i}=\frac{\partial\epsilon(\vec{p})}{\partial p_i}$.
Thus the density of states is proportional to the area of a constant energy surface and inversely proportional to the mean velocity of the particles at the constant energy surface. Thus, flat bands, which correspond to slow particles have a large contribution to the density of states. Steep bands, which are related to fast particles, contribute little to the density of states at a given energy, but they contribute over a large energy range.



In the following we will calculate the density of states of two model systems the free particle with and without mass.

6.2.3 Free particle density of states with mass

The dispersion relation of a particle with mass is

$$\epsilon(p)=\frac{p^2}{2m}$$

The density of states is obtained exploiting the fact that $|\vec{p}|$ is constant on a surface of constant energy

$$\begin{aligned}D(\epsilon)&=\frac{V}{(2\pi\hbar)^3}\oint dA\,\frac{1}{|\nabla_p\epsilon|}=\frac{V}{(2\pi\hbar)^3}\underbrace{4\pi p^2}_{\int dA}\frac{1}{|\vec{p}|/m}\\&\stackrel{|p|=\sqrt{2m\epsilon}}{=}\frac{V}{(2\pi\hbar)^3}4\pi(\sqrt{2m\epsilon})^2\frac{1}{\sqrt{2m\epsilon}/m}\\&=2\pi V\left(\frac{\sqrt{2m}}{2\pi\hbar}\right)^3\sqrt{\epsilon}\end{aligned}\tag{6.10}$$

One should remember that the density of states has a square-root behavior in three dimensions. Note however that two-dimensional particles such as a 2-dimensional electron gas has a radically different density of states. A one-dimensional particle, such as an electron in a one-dimensional conductor has a density of states proportional to $\frac{1}{\sqrt{\epsilon}}$, a particle moving in two dimensions has a density of states which is constant.

This model is important to describe the behavior of states at band edges. For example the electron states of a semiconductor at both sides of the Fermi level are the ones relevant for the electric and thermodynamic properties of a semiconductor. When we approximate the band about the minimum to second order, we obtain an approximate dispersion relation

$$\epsilon(p)=\epsilon_c+\frac{1}{2}(\vec{p}-\vec{p}^0)\mathbf{m}^{*-1}(\vec{p}-\vec{p}^0)$$

where ϵ_c is the conduction band minimum, $\vec{p}^0$ is the momentum of the minimum and $\mathbf{m}^*$ is the effective mass tensor. $\mathbf{m}^*$ is nothing but the inverse of the second derivatives of $\epsilon(\vec{p})$ at the band edge. It is important to note that the effective mass can be a tensor. Furthermore there may be several degenerate minima, so that one obtained different types of conduction electrons.

Similarly we can expand the top of the valence band to quadratic order about its maximum. These are the hole bands. The curvature and thus the effective mass is negative. Since it is a missing electron that conducts, the hole has also a positive charge. The hole is conceptually analogous to the antiparticle of the electron, the positron. It is quite common that concepts from elementary particle physics carry over into solid state physics. In the latter, however, the phenomena can be observed at much lower energies and put into practical use.

Interesting is also the behavior in lower dimensions

$$D^d(\epsilon)=\left(\frac{\sqrt{2mL}}{2\pi\hbar}\right)^d S_d\theta(\epsilon)\epsilon^{\frac{d}{2}-1}=\begin{cases}\frac{\sqrt{2mL}}{2\pi\hbar}2\theta(\epsilon)\epsilon^{-\frac{1}{2}} & \text{for } d=1\\ \left(\frac{\sqrt{2mL}}{2\pi\hbar}\right)^2 2\pi\theta(\epsilon)\epsilon & \text{for } d=2\\ \left(\frac{\sqrt{2mL}}{2\pi\hbar}\right)^3 4\pi\theta(\epsilon)\epsilon^{\frac{1}{2}} & \text{for } d=3\end{cases}\tag{6.11}$$

where $S_d=2d^{\frac{d}{2}}/\Gamma(\frac{1}{2}d)$, that is $S_1=2, S_2=2\pi, S_3=4\pi$, is the surface area of a unit sphere in d dimensions and $\Gamma(x)$ is the Gamma function <(`http://mathworld.wolfram.com/Hypersphere.html`). $\theta(x)$ is the Heaviside step function which vanishes for negative arguments and equals unity for positive arguments.

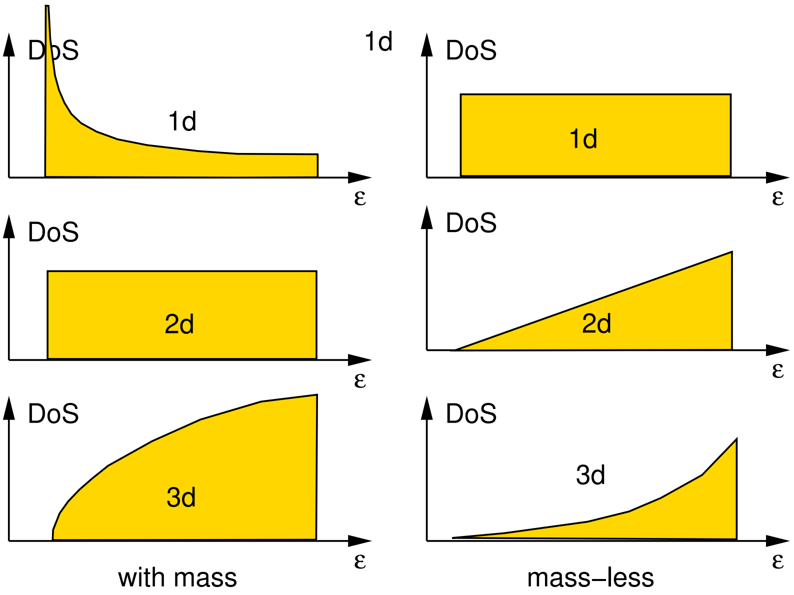


Fig. 6.3: Density of states for particles with and without mass, that is with linear and parabolic dispersion in one, two and three dimensions

6.2.4 Free particle density of states without mass

Mass-less particles are used to describe, for example, the acoustical branch of phonons (lattice vibrations). It can also used to describe electrons in a metal: In thermodynamics only the electrons near the Fermi level are of relevance, so that the band structure $\epsilon(k)$ can be approximated in this region by a linear dispersion relation. Another example where the dispersion relation is linear is light, i.e. photons.

The dispersion relation of a particle without mass is

$$\epsilon(p) = c|\vec{p}|$$

where c is the speed of light, if we discuss photons, or the speed of sound if we discuss phonons, that is lattice vibrations. For metals c is called the Fermi velocity.

The density of states is obtained exploiting the fact that $|\vec{p}|$ is constant on a surface of constant energy

$$D^d(\epsilon) = \left(\frac{L}{2\pi\hbar c}\right)^d S_d \theta(\epsilon)\epsilon^{d-1} = \begin{cases} \frac{L}{2\pi\hbar c} 2\theta(\epsilon) & \text{for } d=1 \\ \left(\frac{L}{2\pi\hbar c}\right)^2 2\pi\theta(\epsilon)\epsilon & \text{for } d=2 \\ \left(\frac{L}{2\pi\hbar c}\right)^3 4\pi\theta(\epsilon)\epsilon^2 & \text{for } d=3 \end{cases} \tag{6.12}$$

6.3 Real and reciprocal lattice

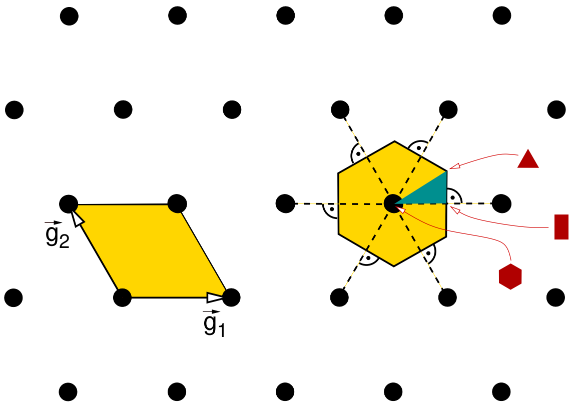


Fig. 6.4: Illustration of the Brillouin zone and high-symmetry points for a plane hexagonal lattice. On the left the unit cell of the reciprocal lattice is shown. On the right the Brillouin zone is shown, which consists of all points that are closer to one lattice point than to any other. In the Brillouin zone there is the irreducible zone, which contains all points that are symmetry inequivalent. The corners of the irreducible zone are high-symmetry points.

Most solids are crystals. This implies that the atoms are located on regular positions. The clearest evidence of this regular arrangement are the crystal faces of gem stones. This arrangement results in a symmetry, the lattice periodicity. Lattice translation symmetry is a discrete translational symmetry.

6.3.1 Real and reciprocal lattice: One-dimensional example

Let us consider a linear chain of hydrogen atoms as the most simple model of a one-dimensionally periodic crystal of hydrogen atoms shown in Fig. 6.5. It is a model system, which is not stable in reality. However it allows to develop the concepts.

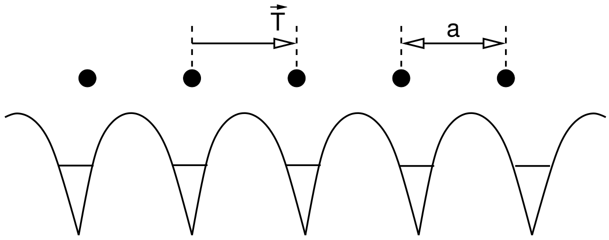


Fig. 6.5: Linear chain of hydrogen atoms. The lattice spacing is a. The balls represent the atoms. The graph sketched below represents the potential as function of position. The primitive lattice vector is $\vec{T}$ which has the length a. The primitive reciprocal lattice vector has the length $2\pi/a$.

The periodicity of the chain implies that for every atom at position $\vec{R}_0$ there are also equivalent atoms at the positions $\vec{R}_n = \vec{R}_0 + n\vec{T}$, where n is any integer. $\vec{T}$ is the **primitive lattice vector**. **General lattice vectors** are $\vec{t}_n = n\vec{T}$. The primitive lattice vector is the smallest possible lattice vector, that fully describes the lattice periodicity. Its length is called the **lattice constant** a.

Not only the atomic positions are regular, but also the potential is **periodic**, that is $v(\vec{r} + \vec{t}_n) = v(\vec{r})$. A periodic potential has a discrete Fourier transform

$$v(\vec{r}) = \sum_n \mathrm{e}^{i\vec{G}_n\vec{r}} V_n$$

where the vectors $\vec{G}_n = n\vec{g}$ are the **general reciprocal lattice vectors**. The **primitive reciprocal lattice vector** is denoted by $\vec{g}$. The length of the primitive reciprocal lattice vector is inversely proportional to the real-space lattice constant a, that is $|\vec{g}| = \frac{2\pi}{a}$

6.3.2 Real and reciprocal lattice in three dimensions

Repeated in section 4.3.1 of ΦSX: *Introduction to Solid State Theory*[14].

Let us now generalize the expressions to three a three-dimensional lattice. Fig. 6.3.2 demonstrates the concepts developed in the following. Corresponding to the three spatial directions there are now three primitive lattice vectors, which we denote by $\vec{T}_1, \vec{T}_2, \vec{T}_3$. Note that in two or three dimensions there is no unique choice of primitive lattice vectors. The three primitive lattice vectors span the **primitive unit cell**.

A general lattice vector $\vec{t}_{i,j,k}$ can be expressed by the **primitive lattice vectors** $\vec{T}_1, \vec{T}_2, \vec{T}_3$ as

$$\vec{t}_{i,j,k} = i\vec{T}_1 + j\vec{T}_2 + k\vec{T}_3 \hat{=} \underbrace{\begin{pmatrix} T_{x,1} & T_{x,2} & T_{x,3} \\ T_{y,1} & T_{y,2} & T_{y,3} \\ T_{z,1} & T_{z,2} & T_{z,3} \end{pmatrix}}_{\boldsymbol{T}} \begin{pmatrix} i \\ j \\ k \end{pmatrix}$$

where i, j, k are arbitrary integers.

It is often convenient to combine the lattice vectors into a 3×3 matrix $\boldsymbol{T}$ as shown above. Often the atomic positions $\vec{R}_n$ are provided in **relative positions** $\vec{X}$ which are defined as

$$\vec{R}_n = \boldsymbol{T}\vec{X}_n \qquad \Leftrightarrow \qquad \vec{X}_n = \boldsymbol{T}^{-1}\vec{R}_n$$

A potential is called periodic with respect to these lattice translations, if $V(\vec{r} + \vec{t}_{i,j,k}) = V(\vec{r})$.

RECIPROCAL LATTICE

The reciprocal lattice is given by those values of the wave vector $\vec{G}$, for which the corresponding plane waves $\mathrm{e}^{i\vec{G}\vec{r}}$ have the same periodicity as the real-space lattice.
The primitive reciprocal-lattice vectors $\vec{g}_n$ for $n = 1, 2, 3$ are defined in three dimensions as

$$\vec{g}_n\vec{T}_m = 2\pi\delta_{n,m} \tag{6.13}$$

Thus, in three dimensions the reciprocal lattice vectors can be obtained as

$$\vec{g}_1 = 2\pi\frac{\vec{T}_2 \times \vec{T}_3}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)} \quad , \quad \vec{g}_2 = 2\pi\frac{\vec{T}_3 \times \vec{T}_1}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)} \quad , \quad \vec{g}_3 = 2\pi\frac{\vec{T}_1 \times \vec{T}_2}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)}$$

It is easily shown that these expressions fulfill the defining equation Eq. 6.13 for the reciprocal lattice vectors. Note that the expressions for the second and third lattice vector are obtained from the first by cyclic commutation of the indices.

Because the vector product is perpendicular to the two vectors forming it, every reciprocal lattice vector is perpendicular to two real-space lattice vectors. This brings us to the physical meaning of the reciprocal lattice vectors: Two real space lattice vectors define a lattice plane. A plane can be defined either by two linearly independent vectors in the plane or alternatively by the plane normal. The reciprocal lattice vectors apparently define lattice planes, because they are plane normals.

Let us now consider the distance Δ_n of two neighboring lattice planes. It is obtained by projecting one vector pointing $\vec{T}_n$ from one plane to the other onto a normalized plane normal $\vec{g}_n/|\vec{g}_n|$.

$$\begin{aligned} \Delta_n &= \vec{T}_n\frac{\vec{g}_n}{|\vec{g}_n|} \overset{\text{Eq. 6.13}}{=} \frac{2\pi}{|\vec{g}_n|} \\ \Rightarrow \qquad |\vec{g}_n| &= \frac{2\pi}{\Delta_n} \end{aligned}$$

PHYSICAL MEANING OF RECIPROCAL LATTICE VECTORS

- The reciprocal lattice vectors are perpendicular to the lattice planes of the real-space lattice.
- The length of the *primitive* reciprocal lattice vectors is 2π divided by the distance of the lattice planes.

We can form a 3×3 matrix $\boldsymbol{g}$ from the three primitive reciprocal lattice vectors.

$$\boldsymbol{g} \hat{\overset{\text{def}}{=}} \begin{pmatrix} g_{x,1} & g_{x,2} & g_{x,3} \\ g_{y,1} & g_{y,2} & g_{y,3} \\ g_{z,1} & g_{z,2} & g_{z,3} \end{pmatrix}$$

The definition Eq. 6.13 of the reciprocal lattice vectors can be expressed in matrix form as

$$\boldsymbol{g}^{\top}\boldsymbol{T} \overset{\text{Eq. 6.13}}{=} 2\pi\mathbf{1} \tag{6.14}$$

Thus the matrix $\boldsymbol{g}$ is obtained as

$$\boldsymbol{g} \overset{\text{Eq. 6.14}}{=} 2\pi\left(\boldsymbol{T}^{-1}\right)^{\top} \tag{6.15}$$

The general reciprocal lattice vectors are

$$\vec{G}_{i,j,k} = i\vec{g}_1 + j\vec{g}_2 + k\vec{g}_3$$

Reciprocal lattice and Fourier transforms

The reciprocal lattice vectors play an important role in the Fourier transform of periodic functions. All plane waves that are periodic with the lattice vectors $\vec{T}_n$ are characterized by a reciprocal lattice vector.

$$
\begin{aligned}
\mathrm{e}^{i\vec{G}(\vec{r}+\vec{T}_n)} &= \mathrm{e}^{i\vec{G}\vec{r}} \\
\Rightarrow \vec{G}(\vec{r}+\vec{T}_n) &= \vec{G}\vec{r}+2\pi m_n \\
\Rightarrow \vec{G}\vec{T}_n &= 2\pi m_n \\
\Rightarrow \vec{G}\boldsymbol{T} &= 2\pi\vec{m} \\
\Rightarrow \vec{G} &= 2\pi\vec{m}\boldsymbol{T}^{-1} = 2\pi\left(\boldsymbol{T}^{-1}\right)^{\top}\vec{m}
\end{aligned}
$$

where m_n are integers. The allowed values for $\vec{G}$ are exactly those of the reciprocal lattice. This was the motivation for the definition Eq. 6.14.

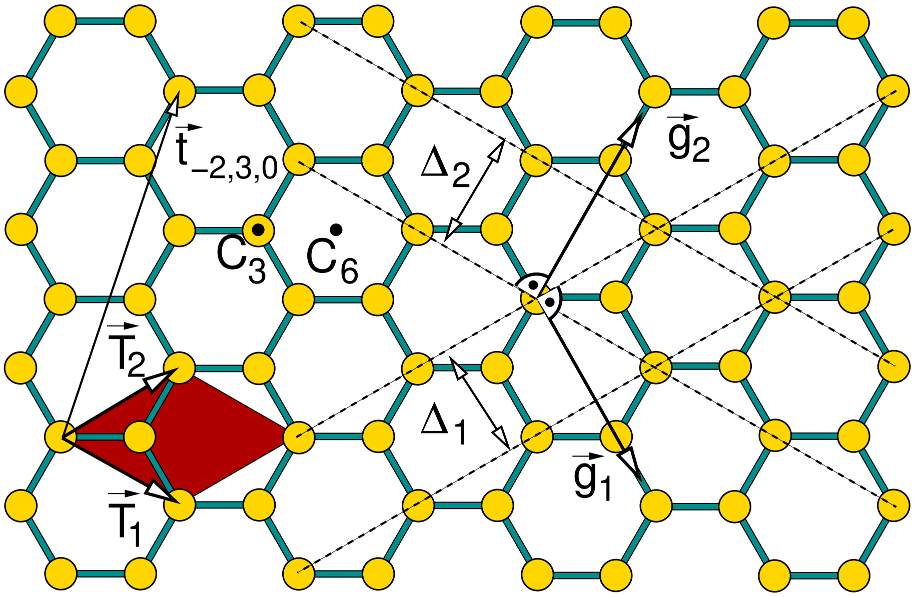


Fig. 6.6: Translational and rotational symmetry of a single graphite sheet and demonstration of reciprocal lattice vectors. Graphite is a layered material, which consists of sheets as the one shown. The yellow balls indicate the carbon atoms and the blue sticks represent the bonds. In graphite these two-dimensional sheets are stacked on top of each other, where every second sheet is shifted such that an atom falls vertically below the point indicated by C_6. Here we only consider the symmetry of a single sheet. The elementary unit cell is indicated by the red parallelogram, which is spanned by the elementary lattice vectors $\vec{T}_1$ and $\vec{T}_2$. The third lattice vector points perpendicular to the sheet towards you. An example for a general lattice vector is $\vec{t}_{3,-2,0} = 3\vec{T}_1 - 2\vec{T}_2 + 0\vec{T}_3$. The lattice planes indicated by the dashed lines. The lattice planes are perpendicular to the sheet. The distance of the lattice planes are indicated by Δ_1 and Δ_2. The elementary reciprocal lattice vectors are $\vec{g}_1$ and $\vec{g}_2$. The third reciprocal lattice vector points perpendicular out of the plane towards you. Note that the reciprocal lattice vectors have the unit “inverse length”. Thus their length is considered irrelevant in this real space figure. The axis through C_3 standing perpendicular to the plane is a three-fold rotation axis of the graphite crystal. The axis through C_6 perpendicular to the plane is a 6-fold rotation axis of the graphite sheet, but only a three-fold rotation axis of the graphite crystal. (Note that a rotation axis for the graphite crystal must be one for both sheets of the crystal). In addition there is a mirror plane lying in the plane. Furthermore there are several mirror planes perpendicular to the plane: One passing through every atom with one of the three bonds lying in the plane and one perpendicular each bond passing through the bond center.

6.4 Bloch Theorem

In order to make the description of an infinte crystal tractable, we exploit translational symmetry to block-diagonalize the Hamiltonian. The theoretical basis is Bloch’s theorem.

BLOCH'S THEOREM

The eigenstates of the Hamiltonian with crystal symmetry $\hat{S}(\vec{t})$, i.e. $[\hat{H}, \hat{S}(\vec{t})]_- = 0$, can be written as a product

$$\psi_{\vec{k},n}(\vec{r}) = \vec{u}_{\vec{k},n}(\vec{r})\mathrm{e}^{i\vec{k}\vec{r}} \tag{6.16}$$

of a periodic function $u_{\vec{k},n}(\vec{r})$ with a plane wave $\mathrm{e}^{i\vec{k}\vec{r}}$. $\hat{S}(\vec{t})$ is a translation operator for the real space translation vector $\vec{t}$.

...or...

The Hamiltonian of a crystal is block diagonal for Bloch states, that is

$$\langle\psi_{\vec{k}}|\hat{H}|\psi_{\vec{k}'}\rangle = 0 \qquad \text{for } \vec{k} \neq \vec{k}'$$

when

$$\psi_k(\vec{r}) = \vec{u}_k(\vec{r})\mathrm{e}^{i\vec{k}\vec{r}} \qquad \text{where } u_k(\vec{r}+\vec{t}) = u_k(\vec{r})$$

The underlying idea of Bloch's theorem is to Block-diagonalize the Hamiltonian by finding the eigenstates of the translation operators.

Let $\hat{S}(\vec{t})$ be the translation operator for a given lattice translation $\vec{t} = \vec{T}_1 i + \vec{T}_2 j + \vec{T}_3 k$.

$$\hat{S}(\vec{t}) = \int d^3r\ |\vec{r}+\vec{t}\rangle\langle\vec{r}| \tag{6.17}$$

Translation operator and canonical momentum

A general theorem is that any translation for an unbounded continuous variable can be expressed by the canonical momentum.

$$\hat{S}(\vec{t}) = \mathrm{e}^{-\frac{i}{\hbar}\hat{\vec{p}}\vec{t}} \tag{6.18}$$

This relation will be used for the Proof of Bloch's theorem. It is discussed to some extend in chapter 9 of ΦSX: Quantum Physics. Let us anyway go quickly over the proof.

Proof of Eq. 6.18:

1. First we define a basis of momentum eigenstates[1]

$$\begin{aligned} \hat{\vec{p}}|\vec{p}\rangle &= |\vec{p}\rangle\vec{p} && \text{eigenvalue equation} \\ \langle\vec{p}|\vec{p}\rangle &= (2\pi\hbar)^3\delta(\vec{p}-\vec{p}') && \text{orthonormality} \\ \hat{1} &= \int \frac{d^3p}{(2\pi\hbar)^3}\,|\vec{p}\rangle\langle\vec{p}| && \text{completeness} \end{aligned}$$

   The real space wave function of the momentum eigenstates are plane waves

$$\langle\vec{r}|\vec{p}\rangle = \mathrm{e}^{\frac{i}{\hbar}\vec{p}\vec{r}} \qquad \text{real space representation}$$

   which provides a transformation between the momentum and position eigenstates.

2. Let us now work out a matrix element that we will need in the following:

$$\langle\vec{r}|\mathrm{e}^{-\frac{i}{\hbar}\hat{\vec{p}}\vec{t}}|\vec{p}\rangle = \underbrace{\langle\vec{r}|\vec{p}\rangle}_{\mathrm{e}^{\frac{i}{\hbar}\vec{p}\vec{r}}}\mathrm{e}^{-\frac{i}{\hbar}\vec{p}\vec{t}} = \mathrm{e}^{\frac{i}{\hbar}\vec{p}(\vec{r}-\vec{t})} = \langle\vec{r}-\vec{t}|\vec{p}\rangle$$

---

[1] There are various definitions for momentum eigenstates in the literature, that differ by their normalization.

3. Now we are ready to bring the operator defined by Eq. 6.18 into the form Eq. 6.17, which is the definition of the translation operator.

$$
\begin{aligned}
\mathrm{e}^{-\frac{i}{\hbar}\hat{\vec{p}}\vec{t}} &= \underbrace{\int d^3r\, |\vec{r}\rangle\langle\vec{r}|}_{\hat{1}}\, \mathrm{e}^{-\frac{i}{\hbar}\hat{\vec{p}}\vec{t}} \underbrace{\int \frac{d^3p}{(2\pi\hbar)^3}\, |\vec{p}\rangle\langle\vec{p}|}_{\hat{1}} \\
&= \int d^3r\, |\vec{r}\rangle \int \frac{d^3p}{(2\pi\hbar)^3}\, \underbrace{\left(\langle\vec{r}|\mathrm{e}^{-\frac{i}{\hbar}\hat{\vec{p}}\vec{t}}|\vec{p}\rangle\right)}_{\langle\vec{r}-\vec{t}|\vec{p}\rangle}\langle\vec{p}| \\
&= \int d^3r\, |\vec{r}\rangle\langle\vec{r}-\vec{t}| \underbrace{\int \frac{d^3p}{(2\pi\hbar)^3}\, |\vec{p}\rangle\langle\vec{p}|}_{\hat{1}} \\
&= \int d^3r\, |\vec{r}+\vec{t}\rangle\langle\vec{r}| \overset{\text{Eq. 6.17}}{=} \hat{S}(t) \qquad\qquad q.e.d
\end{aligned}
$$

Proof of Bloch's theorem

Let us assume that $|\psi\rangle$ is an eigenstate of the lattice translation, i.e.

$$
\underbrace{\langle\vec{r}|\hat{S}(\vec{t})|\psi\rangle}_{\psi(\vec{r}-\vec{t})} \overset{\text{Eq. 6.18}}{=} \underbrace{\langle\vec{r}|\psi\rangle}_{\psi(\vec{r})} \underbrace{\mathrm{e}^{-\frac{i}{\hbar}\vec{p}\vec{t}}}_{\mathrm{e}^{-i\vec{k}\vec{t}}}
$$

Now we multiply this equation with $\mathrm{e}^{-i\vec{k}(\vec{r}-\vec{t})}$

$$
\psi(\vec{r}-\vec{t})\mathrm{e}^{-i\vec{k}(\vec{r}-\vec{t})} = \psi(\vec{r})\mathrm{e}^{-i\vec{k}\vec{r}}
$$

which says that the function u defined as

$$
u(\vec{r}) = \psi(\vec{r})\mathrm{e}^{-i\vec{k}\vec{r}}
$$

is periodic. Thus the eigenstate of the lattice translation is a product of a periodic function $u(\vec{r})$ and a plane wave.

We have shown before that the eigenstates of a Hamiltonian can be expressed a superposition of eigenstates of its symmetry operator. Thus also the eigenstates of the Hamiltonian with lattice symmetry can be written in the form of a Bloch wave, and the wave vector is a quantum number for this system. This finishes the proof of Bloch's theorem.

6.5 Reduced zone scheme

Here we will see that the band structure is periodic in reciprocal space, i.e.

$$
\epsilon_n(\vec{k}) = \epsilon_n(\vec{k}+\vec{G})
$$

and try to understand the implications.

Let us consider an eigenstate of the Hamiltonian, which is written as a Bloch state with a periodic part $u_{\vec{k},n}(\vec{r})$. We expand it by $\mathrm{e}^{i\vec{G}\vec{r}}$ with a reciprocal lattice vector, which itself is periodic.

$$
\Psi_{\vec{k},n}(\vec{r}) = u_{\vec{k},n}(\vec{r})\mathrm{e}^{i\vec{k}\vec{r}} = u_{\vec{k},n}(\vec{r})\mathrm{e}^{-i\vec{G}\vec{r}}\mathrm{e}^{i(\vec{k}+\vec{G})\vec{r}} = u_{\vec{k}+\vec{G},n}(\vec{r})\mathrm{e}^{i(\vec{k}+\vec{G})\vec{r}} = \Psi_{\vec{k}+\vec{G},n}(\vec{r})
$$

We have introduced

$$
u_{\vec{k}+\vec{G},n}(\vec{r}) \overset{\text{def}}{=} u_{\vec{k},n}(\vec{r})\mathrm{e}^{-i\vec{G}\vec{r}}
$$

which is the periodic function corresponding to an eigenstate with a shifted wave vector $\vec{k}+\vec{G}$. Because the wave function is still the same, the eigenvalues for $\vec{k}$ and for $\vec{k}+\vec{G}$ are the same. This shows that the band structure is periodic in reciprocal space.

$$\epsilon_n(\vec{k}) = \epsilon_n(\vec{k}+\vec{G})$$

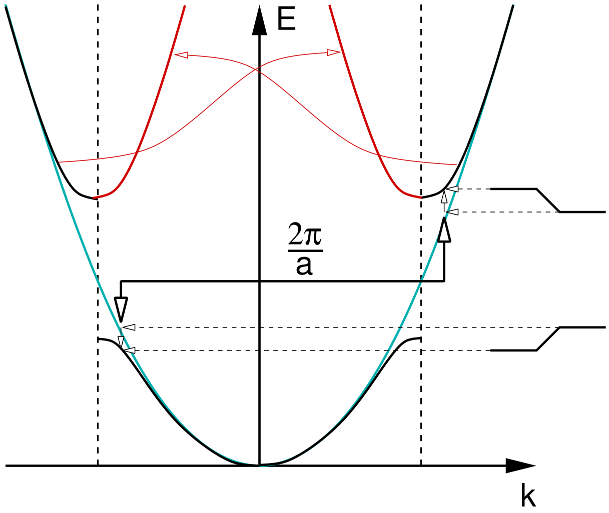


Fig. 6.7: Schematic demonstration how a periodic potential results in a coupling between states with the same wave vector in the reduced zone scheme. If a free particle experiences a potential with periodicity a, states with wave vectors that differ by a reciprocal lattice vector $G_n = \frac{2\pi}{a}n$ couple. The coupling shifts the upper state up and the lower state down, which results in an opening of the gap. As a result of the coupling, the wave vector of a free particle state is no more a good quantum number. However, the wave vector in the reduced zone scheme remains to be a good quantum number.

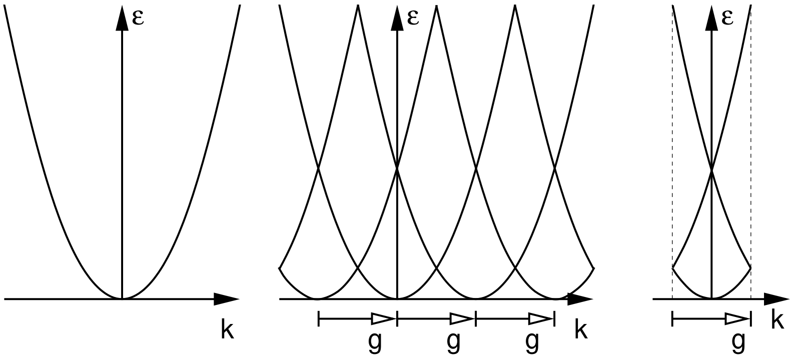


Fig. 6.8: Extended, periodic and reduced zone scheme for a one-dimensional free particle.

The periodicity in reciprocal space is the reason why band structures are not represented in an **extended zone scheme** but in a **reduced zone scheme**.

- in the extended zone scheme the band structure of free electrons would be a single parabola, and would extend to infinity in reciprocal space.

- in the **periodic zone scheme**, the free electron would consists of many parabola centered at the reciprocal lattice points. Thus, the reciprocal zone scheme is periodic with the periodicity of the reciprocal lattice.

- in the reduced zone scheme, only the irreducible part of the periodic zone scheme is shown, that is one single repeat unit. The band structures shown in Fig. 6.1 are in a reduced zone scheme. There are several choices for the repeat units of the reciprocal lattice. One choice is simply the primitive unit cell of the reciprocal lattice. However, the shape of the unit cell does not have the point-group symmetry of the reciprocal lattice. Therefore one instead chooses the **Wigner Seitz cell**[2] of the reciprocal lattice, which is called the **Brillouin zone**.

Opening of gaps at the Brillouin zone boundary

See figure 6.7.

6.6 Bands and orbitals

Repeated in section 4.5 of ΦSX:Introduction to Solid-State Theory[14]

So far we have discussed the band structures as a variation of the dispersion relation of the free electron bands. However this point of view obscures the relation to chemical binding and the local picture of the wave functions.

The band structure can also be constructed starting from local orbitals. This is demonstrated here for the two-dimensional square lattice with atoms having s- and p-orbitals in the plane.

In order to simplify the work, we restrict ourselves to the high symmetry points Γ, X and M of the two-dimensional reciprocal unit cell.

The real space lattice vectors are

$$\vec{T}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} a \qquad \text{and} \qquad \vec{T}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} a$$

The corresponding reciprocal space lattice vectors are

$$\vec{g}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{2\pi}{a} \qquad \text{and} \qquad \vec{g}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{2\pi}{a}$$

The high-symmetry points are

$$k_\Gamma = 0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \frac{2\pi}{a} \quad ; \quad k_X = \frac{1}{2}\vec{g}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{\pi}{a} \qquad \text{and} \qquad k_M = \frac{1}{2}(\vec{g}_1 + \vec{g}_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{\pi}{a}$$

The first step is to construct Bloch waves out of the atomic orbitals, by multiplying them with $e^{i\vec{k}\vec{t}}$, where $\vec{t}$ is the lattice translation of the atom relative to the original unit cell. This is illustrated in Fig. 6.9 on page 97.

- At the Γ, the phase factor for a lattice translation of the orbital is one. Thus we repeat the orbital from the central unit cell in all other unit cells.

---

[2] The Wigner Seitz cell of a lattice consists of all points that are closer to the origin than to any other lattice point. Thus, it is enclosed by planes perpendicular to the lattice vectors cutting the lattice vector in half.

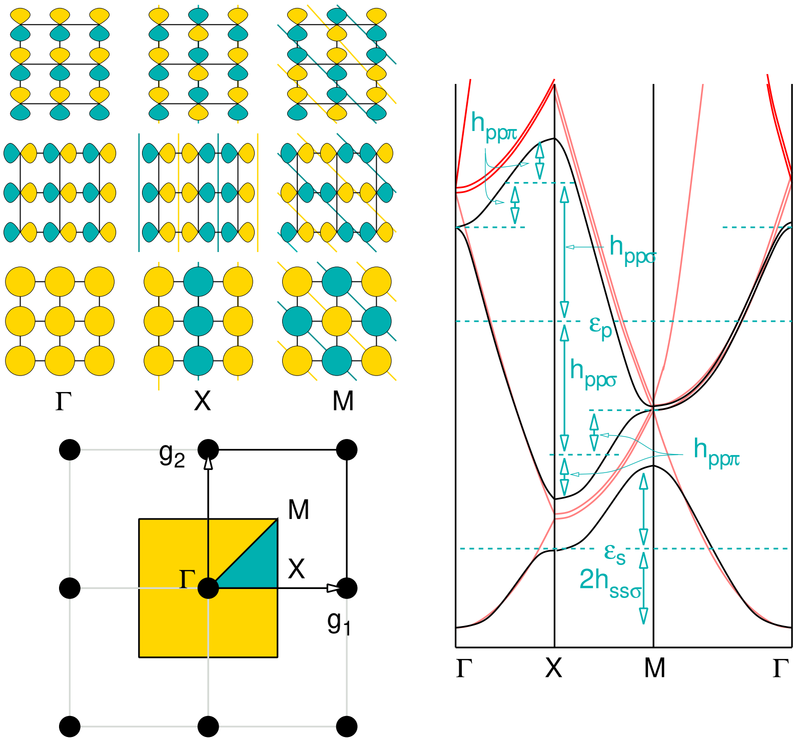


Fig. 6.9: Illustration of a band structure in real and reciprocal space for a planar square lattice. Bottom left: The reciprocal lattice is spanned by the reciprocal lattice vectors $\vec{g}_1$ and $\vec{g}_2$.The yellow square is the corresponding Brillouin zone. The green inset is the Brillouin zone, whose corners are the high symmetry points Γ, X and M. On the top left the corresponding basis functions made from one s-orbital and two p-orbitals in a Bloch-basis are shown. These states are also eigenstates of the Hamiltonian, because the coupling vanishes due to point group symmetry at the high-symmetry points. On the top left a schematic band structure is shown and compared to a free-electron band (red).

- At the X point, the phase factor for a lattice translation along the first lattice vector is $\mathrm{e}^{i\vec{k}_x\vec{T}_1}=\mathrm{e}^{i\pi}=-1$ and the phase factor for the lattice translation along the second lattice vector $\vec{T}_2$ is $\mathrm{e}^{i\vec{k}_x\vec{T}_2}=+1$. Thus if we go to the right the sign of the orbital alternates, while in the vertical direction the orbital remains the same.
- At the M point, the phase factor alternates in each lattice direction, resulting in a checkerboard pattern.

Next we need to set up the Hamilton matrix elements for each k-point, using the Slater-Koster parameters. For a back-on-the envelope estimate the nearest neighbor matrix elements will be sufficient. The matrix elements are firstly ϵ_s and ϵ_p, the energies of the orbitals without any hybridization. Secondly we need the hopping matrix elements $t_{ss\sigma}$, $t_{pp\sigma}$, $t_{pp\pi}$. In principle we would also need $t_{sp\sigma}$, but that matrix element will not be needed in our example.

In our simple example, the Bloch states are already eigenstates of the Hamiltonian so that we only need to calculate their energy. We obtain

Γ	X	M
$\epsilon_p+t_{pp\sigma}-t_{pp\pi}$	$\epsilon_p+t_{pp\sigma}+t_{pp\pi}$	$\epsilon_p-t_{pp\sigma}+t_{pp\pi}$
$\epsilon_p+t_{pp\sigma}-t_{pp\pi}$	$\epsilon_p-t_{pp\sigma}-t_{pp\pi}$	$\epsilon_p-t_{pp\sigma}+t_{pp\pi}$
$\epsilon_s+2t_{ss\sigma}$	ϵ_s	$\epsilon_s+2t_{ss\sigma}$

If there is a band structure available, we can estimate the Hamilton matrix elements from that. Otherwise we just make an educated guess.

6.7 Calculating band structures in the tight-binding model

Repeated in section 4.6 of ΦSX:Introduction to Solid-State Theory[14]

Now we wish to calculate the band structure. Let us begin with a tight-binding basisset with the usual assumption that the tight-binding orbitals shall be orthogonal. The tight-binding orbitals shall be represented by kets $|\vec{t},\alpha\rangle$, where $\vec{t}$ denotes a discrete lattice-translation vector and α denotes the type of the orbital such as an $s, p_x, p_y, p_z, \dots$ orbital on a specific site in the unit cell.

A Bloch wave can be represented as

$$\begin{aligned}
|\Psi_{\vec{k},n}\rangle &= \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\langle\vec{t},\alpha|\Psi_{\vec{k},n}\rangle \\
&\overset{\text{Eq. 6.16}}{=} \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\langle\vec{t},\alpha| \int d^3r\; |\vec{r}\rangle \mathrm{e}^{i\vec{k}\vec{r}}\langle\vec{r}|u_{\vec{k},n}\rangle \\
&= \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{t},\alpha| \int d^3r\; |\vec{r}\rangle \mathrm{e}^{i\vec{k}(\vec{r}-\vec{t})}\langle\vec{r}|u_{\vec{k},n}\rangle \\
&\overset{u(\vec{r})=u(\vec{r}-\vec{t})}{=} \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{t},\alpha| \int d^3r\; |\vec{r}\rangle \mathrm{e}^{i\vec{k}(\vec{r}-\vec{t})}\langle\vec{r}-\vec{t}|u_{\vec{k},n}\rangle \\
&\overset{\vec{r}\to\vec{r}+\vec{t}}{=} \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}} \int d^3r\; \underbrace{\langle\vec{t},\alpha|\vec{r}+\vec{t}\rangle}_{\langle\vec{0},\alpha|\vec{r}\rangle} \underbrace{\mathrm{e}^{i\vec{k}\vec{r}}\langle\vec{r}|u_{\vec{k},n}\rangle}_{\langle\vec{r}|\Psi_{\vec{k},n}\rangle} \\
&= \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\psi_{\vec{k},n}\rangle \qquad (6.19)
\end{aligned}$$

This shows us how we can represent the entire wave function with a vector $\vec{c}_n$ with elements $c_{\alpha,n} = \langle\vec{0},\alpha|\Psi_{k,n}\rangle$, which has the dimension of the number of orbitals in the elementary unit cell.

Our next goal is to find an equation for this vector. We will use the matrix elements of the Hamilton operator in the basis of tight-binding orbitals

$$\langle\vec{t},\alpha|\hat{H}|\vec{t}',\beta\rangle = h_{\vec{t},\alpha,\vec{t}',\beta}$$

Let us insert the ansatz Eq. 6.19 into the Schrödinger equation

$$\begin{aligned}
&\hat{H}|\Psi_{k,n}\rangle = |\Psi_{k,n}\rangle\epsilon_{k,n} \\
\overset{Eq.\ 6.19}{\Rightarrow}\quad & \sum_{\vec{t},\alpha} \hat{H}|\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle = \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle\epsilon_{k,n} \\
\overset{\langle\vec{0},\beta|}{\Rightarrow}\quad & \sum_{\vec{t},\alpha} \underbrace{\langle\vec{0},\beta|\hat{H}|\vec{t},\alpha\rangle}_{h_{\vec{0},\beta,\vec{t},\alpha}} \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle = \sum_{\vec{t},\alpha} \underbrace{\langle\vec{0},\beta|\vec{t},\alpha\rangle}_{\delta_{\vec{0},\vec{t}}\delta_{\alpha,\beta}} \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle\epsilon_{k,n} \\
\Rightarrow\quad & \underbrace{\left[\sum_{\vec{t},\alpha} h_{\vec{0},\beta,\vec{t},\alpha}\mathrm{e}^{i\vec{k}\vec{t}}\right]}_{h_{\alpha,\beta}(\vec{k})} \underbrace{\langle\vec{0},\alpha|\Psi_{k,n}\rangle}_{c_{\alpha,n}(\vec{k})} = \underbrace{\langle\vec{0},\beta|\Psi_{k,n}\rangle}_{c_{\beta,n}(\vec{k})}\epsilon_{k,n} \\
\Rightarrow\quad & \boldsymbol{h}(\vec{k})\vec{c}_n(\vec{k}) = \vec{c}_n(\vec{k})\epsilon_n(\vec{k})
\end{aligned}$$

Thus we obtain

SCHRÖDINGER EQUATION IN K-SPACE

$$\boldsymbol{h}(\vec{k})\vec{c}_n(\vec{k}) = \vec{c}_n(\vec{k})\epsilon_n(\vec{k}) \tag{6.20}$$

with $c_{\alpha,n}(\vec{k}) \stackrel{\text{def}}{=} \langle\vec{0},\beta|\Psi_{\vec{k},n}\rangle$, respectively $|\Psi_{\vec{k},n}\rangle \stackrel{\text{Eq. 6.19}}{=} \sum_{\vec{t}\alpha,n} |\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}} c_{\alpha,n}(\vec{k})$

and

$$h_{\alpha,\beta}(\vec{k}) \stackrel{\text{def}}{=} \sum_{\vec{t},\alpha} h_{\vec{0},\beta,\vec{t},\alpha} \mathrm{e}^{i\vec{k}\vec{t}} \tag{6.21}$$

Thus we obtained a k-dependent eigenvalue equation in matrix form with a finite and usually small dimension. This problem can be broken down into smaller subproblems, if we exploit the symmetry arguments that we used earlier for the molecules. Symmetry can be used especially at the high-symmetry points and lines in reciprocal space, where the band structure is usually represented.

Worked example

Let us consider the example of a linear chain of s-orbitals. The infinite Hamiltonian has the form

$$\mathbf{h} = \begin{pmatrix}
 & \vdots & \vdots & \vdots & & & & & \\
\dots & \epsilon_s & h_{ss\sigma} & 0 & \dots & & & & \\
\dots & h_{ss\sigma} & \epsilon_s & h_{ss\sigma} & 0 & \dots & & & \\
\dots & 0 & h_{ss\sigma} & \epsilon_s & h_{ss\sigma} & 0 & \dots & & \\
 & \dots & 0 & h_{ss\sigma} & \epsilon_s & h_{ss\sigma} & 0 & \dots & \\
 & & \dots & 0 & h_{ss\sigma} & \epsilon_s & h_{ss\sigma} & \dots & \\
 & & & \vdots & \vdots & \vdots & \vdots & &
\end{pmatrix}$$

The Hamilton-matrix elements can also be written as

$$h_{i,j} = \epsilon_s \delta_{i,j} + h_{ss\sigma}\left(\delta_{i,j+1} + \delta_{i,j-1}\right)$$

where i and j refer to the lattice translations and the orbital index has been suppressed, as there is only one orbital.

When we transform the Hamiltonian into Bloch representation using Eq. 6.21, we obtain the k-dependent Hamiltonian with $\vec{t}_j = \Delta j$

$$\begin{aligned}
h(\vec{k}) &\stackrel{\text{Eq. 6.21}}{=} \sum_j h(0,j)\mathrm{e}^{ik\Delta j} = +\underbrace{h_{ss\sigma}\mathrm{e}^{-ik\Delta}}_{j=-1} + \underbrace{\epsilon_s}_{j=0} + \underbrace{h_{ss\sigma}\mathrm{e}^{ik\Delta}}_{j=1} \\
&= \epsilon_s + 2h_{ss\sigma}\cos(k\Delta)
\end{aligned}$$

In this simple case with one orbital per unit cell, the Hamiltonian is only a number for each k. Thus the eigenvalue is

$$\epsilon(k) = \epsilon_s + 2h_{ss\sigma}\cos(k\Delta)$$

Note that $h_{ss\sigma}$ is negative, because it is approximately equal to the overlap matrix element multiplied with the potential in the bond-center. The potential is negative.

Worked example: square two-dimensional lattice

The first step is to set up the hamilton matrix elements in real space. For this purpose it is useful to start with a matrix as shown in 6.10, which shows all the orbitals that interact with each other. In each row we take into account all orbitals that have a non-zero hopping matrix element with a single orbital in the first unit cell.

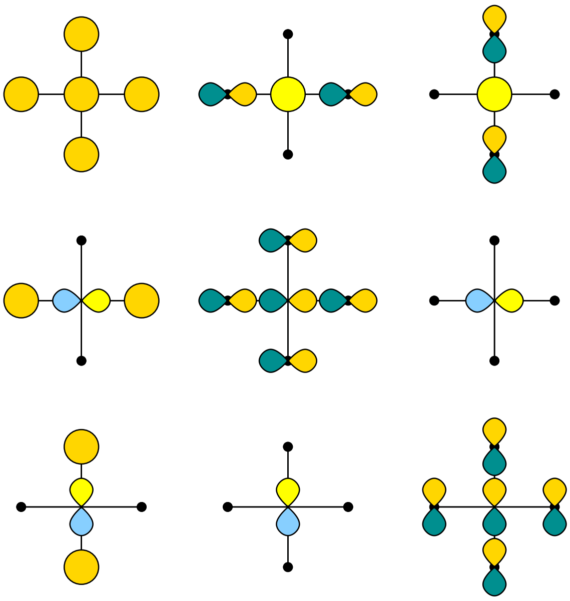


Fig. 6.10: Real space Hamilton matrix elements of a two-dimensional square lattice of atoms with s and p-orbitals. Only the orbitals lying in the plane of the lattice are shown. Each orbitals except for the light-colored ones stand for one Hamilton matrix element between the orbital in the center and that orbital.

$$
\begin{aligned}
\boldsymbol{h}_{(0,0)} &= \begin{pmatrix} \epsilon_s & 0 & 0 \\ 0 & \epsilon_p & 0 \\ 0 & 0 & \epsilon_p \end{pmatrix} \\
\boldsymbol{h}_{(1,0)} &= \begin{pmatrix} t_{ss\sigma} & -t_{sp\sigma} & 0 \\ t_{sp\sigma} & -t_{pp\sigma} & 0 \\ 0 & 0 & t_{pp\pi} \end{pmatrix} \\
\boldsymbol{h}_{(-1,0)} &= \begin{pmatrix} t_{ss\sigma} & t_{sp\sigma} & 0 \\ -t_{sp\sigma} & -t_{pp\sigma} & 0 \\ 0 & 0 & t_{pp\pi} \end{pmatrix} = \boldsymbol{h}^{\dagger}_{(1,0)} \\
\boldsymbol{h}_{(0,1)} &= \begin{pmatrix} t_{ss\sigma} & 0 & -t_{sp\sigma} \\ 0 & t_{pp\pi} & \\ t_{sp\sigma} & 0 & t_{pp\sigma} \end{pmatrix} \\
\boldsymbol{h}_{(0,-1)} &= \begin{pmatrix} t_{ss\sigma} & 0 & -t_{sp\sigma} \\ 0 & t_{pp\pi} & \\ t_{sp\sigma} & 0 & t_{pp\sigma} \end{pmatrix}
\end{aligned}
\tag{6.22}
$$

Now we can set up our hamiltonian via

$$
\begin{aligned}
h_{\alpha,\beta}(\vec{k}) &= h_{\alpha,\beta,\vec{t}}\mathrm{e}^{i\vec{k}\vec{t}} \qquad (6.23)\\
h_{ss}(\vec{k}) &= \epsilon_s \mathrm{e}^{i(k_x\cdot 0a+k_y\cdot 0a)} + t_{ss\sigma}\mathrm{e}^{i(k_x a+k_y\cdot 0a)} + t_{ss\sigma}\mathrm{e}^{i(k_x(-a)+k_y\cdot 0a)} + t_{ss\sigma}\mathrm{e}^{i(k_x\cdot 0a+k_y\cdot a)} + t_{ss\sigma}\mathrm{e}^{i(k_x\cdot 0a+k_y\cdot(-a))}\\
&= \epsilon_s + 2t_{ss\sigma}\Big(\cos(k_x a) + \cos(k_y a)\Big)\\
h_{sp_x}(\vec{k}) &= -t_{sp\sigma}\mathrm{e}^{ik_x a} + t_{sp\sigma}\mathrm{e}^{-ik_x a} = -2it_{sp\sigma}\sin(k_x a)\\
h_{sp_y}(\vec{k}) &= -t_{sp\sigma}\mathrm{e}^{ik_y a} + t_{sp\sigma}\mathrm{e}^{-ik_y a} = -2it_{sp\sigma}\sin(k_y a)\\
h_{p_x s}(\vec{k}) &= t_{sp\sigma}\mathrm{e}^{ik_x a} - t_{sp\sigma}\mathrm{e}^{-ik_x a} = 2it_{sp\sigma}\sin(k_x a)\\
h_{p_x p_x}(\vec{k}) &= \epsilon_p - 2t_{pp\sigma}\cos(k_x a) + 2t_{pp\pi}\cos(k_y a)\\
h_{p_x p_y}(\vec{k}) &= 0\\
h_{p_y s}(\vec{k}) &= t_{sp\sigma}\mathrm{e}^{ik_y a} - t_{sp\sigma}\mathrm{e}^{-ik_y a} = 2it_{sp\sigma}\sin(k_y a)\\
h_{p_y p_x}(\vec{k}) &= 0\\
h_{p_y p_y}(\vec{k}) &= \epsilon_p - 2t_{pp\sigma}\cos(k_y a) + 2t_{pp\pi}\cos(k_x a) \qquad (6.24)
\end{aligned}
$$

In order to draw a band structure along a line you simply evaluate the matrix for the k-points along the line and diagonalize the matrix for each point.

Chapter 7

From atoms to solids

In this chapter I want to describe how one can construct an approximate density of states for materials. The density of states in turn already provides insight into many of the electronic properties of a material. A rough density of states can already be constructed without the knowledge of the atomic structure of a material. Much of it depends only on the composition.

7.1 The Atom

7.1.1 The generalized hydrogen atom

When we think of an atom the first thought is the generalized hydrogen atom. A generalized hydrogen atom is an atom with **non-interacting electrons** and a nuclear charge $+eZ$. The Schrödinger equation of the generalized hydrogen atom can be solved analytically.

The energy levels of a generalized hydrogen atom are

$$\epsilon_{n,\ell,m,s} = -\frac{m_e e^4}{(4\pi\epsilon_0\hbar)^2}\frac{Z^2}{2n^2}$$

The states are characterized by the quantum numbers

- the principal quantum number n
- the angular momentum quantum number $\ell \in \{0,\ldots,n-1\}$. The main angular momenta are als indicated by the letters s, p, d, f for $\ell = 0, 1, 2, 3$.
- the magnetic quantum number $m \in \{-\ell,\ldots,\ell\}$
- the spin quantum number $s \in \{\uparrow,\downarrow\}$.

Many of these quantum numbers are degenerate.

- The degeneracy of different angular momentum quantum numbers ℓ is a consequence of the Z/r potential, which has a dynamical symmetry that conserves the so-called **Laplace-Runge-Lenz vector**[1]. This degeneracy is lifted as soon as the potential obtains a shape different from the $1/r$ shape.
- The degeneracy of the magnetic quantum number is a consequence of the spherical symmetry of the atom.

---

[1] See ΦSX: Quantum mechanics, Appendix V *"ℓ degeneracy of the hydrogen atom"*.

- The degeneracy of the spin eigenvalues is there simply because it does not enter the Hamiltonian.

The generalized hydrogen atom has the following multiplets

n	degeneracy	valence configuration
1	2	2s
2	8	2s+6p
3	18	2s+6p+10d
4	32	2s+6p+10d+14f
5	50	2s+6p+10d+14f+18g
6	72	2s+6p+10d+14f+18g+22h
7	88	2s+6p+10d+14f+18g+22h+16i

7.1.2 Lifting the ℓ-degeneracy: Aufbau principle

The generalized hydrogen atom assumes that the electrons do not interact with each other. If we consider the Coulomb repulsion between the electrons, each electron effectively sees not only the nucleus, but the entire core including the inner electrons.

As a result, the outer electrons experience a potential of the form

$$v_{eff}(r)=\frac{-Z_{eff}e^2}{4\pi\epsilon_0 r}$$

where the effective atomic number $Z_{eff}=Z-N_{core}$ is given by the charge of the nucleus and those of the inner electrons, the core electrons.

Quite a few observations can be rationalized by the simple model, that the Coulomb potential of the nucleus is screened by the inner electrons.

- **Screening of the nuclear Coulomb attraction:** The bare Coulomb potential is too attractive and would place the valence electrons at much to low energy. If the atomic number is reduced to the effective atomic number, the valence electrons of all atoms are placed in a similar energy region. This explains why atoms in the same group behave similar, and why the period they belong to is secondary.

- **Lifting the ℓ degeneracy:** The observation is that

  - for a given atom, the s-electrons always lie below the p-electrons with the same principal quantum number and that
  - the 3d-electrons become occupied only in the 4-th period instead of the 3rd period and
  - that the 4d-electrons become occupied only in the 6-th period instead of the 4th period

  This effect can be traced to the contraction of the core orbitals.

  The "accidental" ℓ degeneracy of the generalized hydrogen atom can be described as a balance between zentrifugal force and core repulsion, that both act against the nuclear Coulomb attraction.

  - The zentrifugal force puts a penalty onto states with larger angular momentum. A higher angular momentum implies a stronger centrifugal force, that pushes electrons away from then nucleus. [2]

---

[2] The concept of centrifugal force can be seen in the radial Schrödinger equation for the radial $R(r)$ part of the

- **Core repulsion**, on the other hand, puts a penalty to states with lower angular momentum. Core repulsion only acts between electrons with the same angular momentum[3]. Because there are more core shells with lower angular momentum, the core repulsion acts stronger on the low-angular momentum states.

In the real atom the subtle balance between zentrifugal force and core repulsion is broken, because the inner electrons are more contracted so that the core repulsion is weakened. The reduced core repulsion stabilizes[4] the states with lower angular momentum.

The core electrons are contracted because there are less inner electron-shells screening the nuclear attraction of the core shells than of the valence electrons.

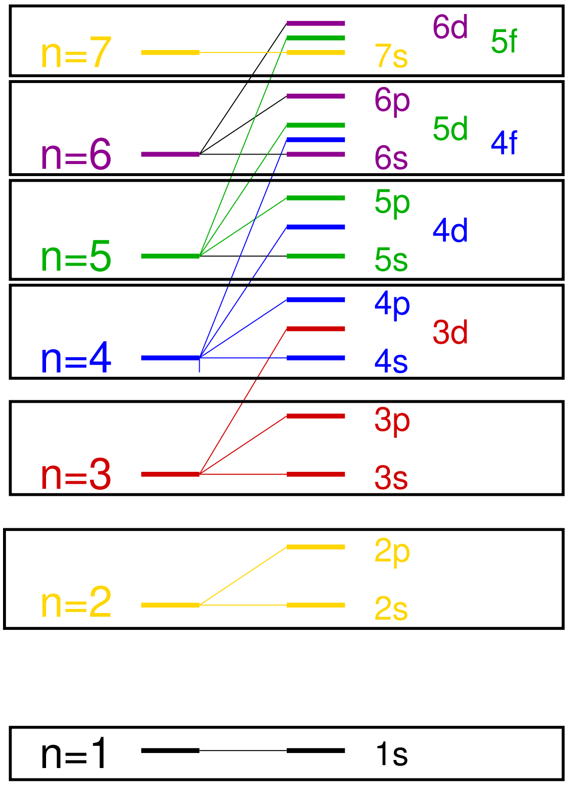


Fig. 7.1: Aufbau principle. Shown is a hypothetical energy diagram constructed such that the occupation from bottom up directly leads to the periodic table. (I have drawn the figure such that it suggests an upward shift of shells with higher angular momentum. This may be misleading as we trace the origin to a stabilization of shells with lower angular momentum.)

---

wave function $\Psi(\vec{r}) = R(|\vec{r}|)Y_{\ell,m}(\vec{r})$

$$\left(-\frac{\hbar^2}{2m}\partial_r^2 + \underbrace{\frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}}_{v_{zf}(r)} - \frac{Ze^2}{2m_0 r} - \epsilon_n\right) rR(\vec{r}) = 0$$

The radial part experiences an additional ℓ-dependent potential that pushes the electrons away from the nucleus.

[3] Core repulsion is due to the condition that wave functions are orthogonal. This condition in turn is a consequence of the Pauli principle. Because states with different angular momentum are orthogonal by their angular motion, there is no additional effect from core orthogonalization. Core repulsion is a variant of Pauli repulsion.

[4] shifts the energy levels down in energy

Radial extent is determined by principal qantum number

The radial extent of the orbitals is, surprisingly, rather similar for orbitals with the same principal quantum number *n*. A consequence is that the 4s and 4p orbitals in a transition metal behave almost like free electrons, while the 3d-orbitals are core-like and interact only weakly with neighboring atoms.

This explains that the d-electrons of transition metal elements form weak bonds that can be broken easily. This makes them ideal for catalytic behavior. A catalyst must form a bond with the transition state of a reaction, thus lowering the reaction barrier, but this bond must also be able to break again, so that the catalyst can return into its original form. The f-electrons behave like true core electrons, with the exception that the f-shell is only partially filled.

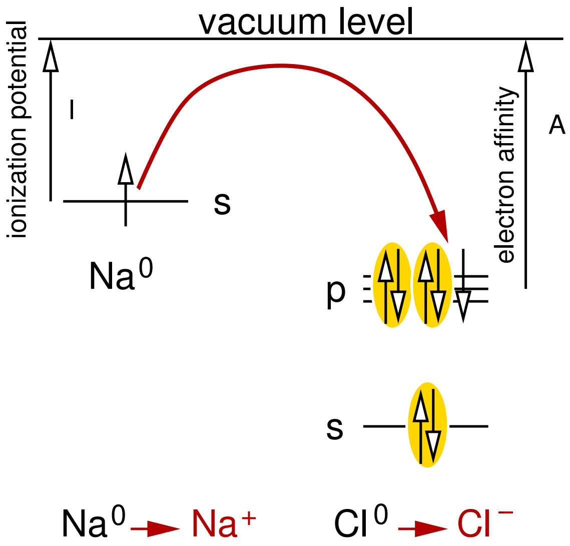
The crossing of 3d and 4s orbitals for increasing atomic number is evident from the periodic table. However, in compounds it is much less evident. In a compound, transition metals normally loose their *s*-electrons. Because the *s*-electrons are located far from the nucleus, they experience the Pauli repulsion from the electrons of any neighboring atoms. This shifts their electron levels up, so that the *s*-electrons are transferred into the *d*-shell. We could say that the Pauli repulsion by the core electrons is compensated by that of the neighboring atoms.

Even when the metals are in their elemental form, these metals can be considered cations immersed into a free electron gas formed by the s-electrons.

7.2 From atoms to ions

7.2.1 Ionization potential and electron affinity

Even before atoms are brought together, they may already exchange electrons. This is the onset for the formation of an ionic bond. My experience is that this should be the first step to rationalize chemical bonding in compounds made from different atoms.



The energy to move an electron to the **vacuum level** is the **ionization potential**. The energy to move an electron from the vacuum level to an atom is called the **electron affinity**.

Thus the energy gained by transferring electrons is the difference between the ionization potential and the electron affinity.

$$\Delta E_1 = A - I$$

Once the ions are formed, they attract each other electrically, and additional energy

$$\Delta E_2 = \frac{Q_1 Q_2}{4\pi\epsilon_0 |\vec{R}_1 - \vec{R}_2|}$$

is gained by bringing them closer together.

The third effect playing a role is the onsite Coulomb repulsion of the electrons. Adding a second electron to an atom is more difficult than adding the first, because the two additional electrons repell each other. Therefore, the electron gain for transferring electrons becomes weaker with each electron already transferred.

The direction of the charge transfer can be read from the **electronegativities**. Mulliken's electronegativity[15] is proporational to the mean value of electron affinity and ionization potential. Electrons are transferred from the electropositive (less electronegative) atoms to the more electronegative atoms. Therefore the first step that we consider when bringing atoms into a material, is to transfer electrons, so that the electropositive atoms give electrons to electronegative atoms. Electropositive atoms are predominantly to the left of the periodic table and electronegative atoms are on the right.

The limit for the electron transfer is reached when an electron shell is completely full or empty.

7.2.2 Oxidation states and electron count

In my experience, the best way to rationalize the electronic structure of a solid material is to use the fully ionic model as starting point.

This electron transfer is not always unique. If it is not obvious how the electrons are transferred, one can consider both choices or select one of them. If it is not obvious how to choose, it is also not important.

Let us do some examples. Use the periodic table to identify the number and type of available valence electrons.

- $NaCl \rightarrow Na^+Cl^-$
- $MgO \rightarrow Mg^{2+}O^{2-}$
- $SiO_2 \rightarrow Si^{4+}O_2^{2-}$
- $SrTiO_3 \rightarrow Sr^{2+}Ti^{4+}O_3^{2-}$
- $LaAlO_3 \rightarrow La^{3+}Al^{3+}O_3^{2-}$
- CaMnO3$\rightarrow Ca^{2+}Mn^{4+}O_3^{2-}$. The Mn atom keeps three some of its seven d-electrons. The number of d-electrons left on Mn is determined by the number of electrons that can be passed to the oxygen atoms
- $CO_2 \rightarrow C^{4+}O_2^{2-}$
- $GaAs \rightarrow Ga^{3+}As^{3-}$. GaAs is a $III - V$ compound. It belongs to the same structural class as silicon, but instead of four-valent electrons its consists of atoms with three and with five valence electrons.

Interesting is the case at surfaces or interfaces. Take as an example the interface between $SrTiO_3$ and $LaAlO_3$. $SrTiO_3$ is a perovskite which consists of alternating planes of TiO_2 and SrO. $LaAlO_3$ has the same structure wit planes of AlO_2 and LaO. A the interface a 2-dimensional electron gas is formed because the charges at the interface cannot becompensated.

Another case is the polar (111) surface of an oxide such as Al_2O_3. The oxide has alternating layers of Al and O. At the oxygen terminated surface there is a net charge. This charge has to enter the Al-states, which is unfavorable. In humid air, it can also react with water so that the oxygen atoms are converted into hydroxyl groups having less net charge.

A common misconception is that the ionic construction reflects the real charge distribution. There are various definitions of the charge of an atom. Here we use the concept of an oxidation stateWhat we use here are the oxidation states. The oxidation state an extremely valuable tool to predict the position of band gaps and of the stability of a material.

The **oxidation state** is a hypothetical charge that an atom would have if all bonds would be fully ionic. (See IUPAC definition on p. 108)

The oxidation state is based on the convention that an electron is fully attributed to the more electron acceptor. In reality, however, this orbital is a superposition of orbitals of electron donor and electron acceptor. Therefore some of the charge distribution remains at the electron donor. This explains the seeming contradiction between charge distribution and formal charges.

Even for an ionic compound, the charge density does not deviate much from that of overlapping atomic charge densities. The formation of an ionic bond does not lead to a large rearrangement of the charge distribution because of the following: The electrons of an electropositive atom are loosely bound and therefore located at large distance from the nucleus. In a compound, the nearest neighbors are located at the same distance, so that little charge rearrangement is required to transfer an electron.

OXIDATION STATE

“A measure of the degree of oxidation of an atom in a substance. It is defined as the charge an atom might be imagined to have when electrons are counted according to an agreed-upon set of rules:

1. the oxidation state of a free element (uncombined element) is zero;
2. for a simple (monatomic) ion, the oxidation state is equal to the net charge on the ion;
3. hydrogen has an oxidation state of 1 and oxygen has an oxidation state of -2 when they are present in most compounds. (Exceptions to this are that hydrogen has an oxidation state of -1 in hydrides of active metals, e.g. LiH, and oxygen has an oxidation state of -1 in peroxides, e.g. H_2O_2;
4. the algebraic sum of oxidation states of all atoms in a neutral molecule must be zero, while in ions the algebraic sum of the oxidation states of the constituent atoms must be equal to the charge on the ion. For example, the oxidation states of sulfur in H_2S, S_8 (elementary sulfur), SO_2, SO_3, and H_2SO_4 are, respectively: -2, 0, +4, +6 and +6. The higher the oxidation state of a given atom, the greater is its degree of oxidation; the lower the oxidation state, the greater is its degree of reduction.”

Source IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://goldbook.iupac.org (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi:10.1351/goldbook.

Chapter 8

Instabilities

We already have seen that if an occupied and an unoccupied state is brought together, a bond is formed. As a result the occupied state is lowered and the unoccupied state is raised in energy. Thus a gap opens up. This effect can lead to symmetry-breaking distortions. We are used that nature sort of likes symmetries. Therefore it is important to see when symmetry is broken.

Let us consider the Hamiltonian with two states

$$\mathbf{H} = \begin{pmatrix} \bar\epsilon_1 & t \\ t & \bar\epsilon_2 \end{pmatrix}$$

The eigenvalues are according to Eq. 2.8

$$\epsilon_\pm = \frac{\bar\epsilon_1 + \bar\epsilon_2}{2} \pm \sqrt{(\frac{\bar\epsilon_1 - \bar\epsilon_2}{2})^2 + t^2} \tag{8.1}$$

Let us make a model where the energy is given by the sum of one-particle energies and an additional term, that describes a symmetry lowering coordinate, that is coupled to the parameter t.

To make things simpler we consider two electrons. Our total energy for the model is

$$E = 2\frac{\bar\epsilon_1 + \bar\epsilon_2}{2} - 2\sqrt{(\frac{\bar\epsilon_1 - \bar\epsilon_2}{2})^2 + t^2} + \frac{1}{2}ct^2$$

If the two energies $\bar\epsilon_1$ and $\bar\epsilon_2$ are non-degenerate the total energy has the approximate form

$$E(t) = 2\bar\epsilon_1 - 2\frac{t^2}{|\epsilon_2 - \epsilon_1|} + \frac{1}{2}ct^2 + O(t^4)$$

Thus the system is stable if $c > \frac{4}{|\epsilon_2 - \epsilon_1|}$ and unstable otherwise. It is clear that the system is likely to develop an instability of the two energies are close together.

For degenerate energies the situation changes drastically. In that case our energy has the form

$$E(t) = 2\bar\epsilon_1 - |t| + \frac{1}{2}ct^2 + O(t^4)$$

Thus an instability is bound to develop for degenerate levels, no matter how strong the force-constant c is.

What is needed is a symmetry, so that the energy of the non-electronic degrees of freedom form a minimum with a zero derivative at the symmetry center. Secondly we need states that are degenerate in the symmetry center. Thirdly we need an occupation so that a degenerate multiplet is only partially occupied.

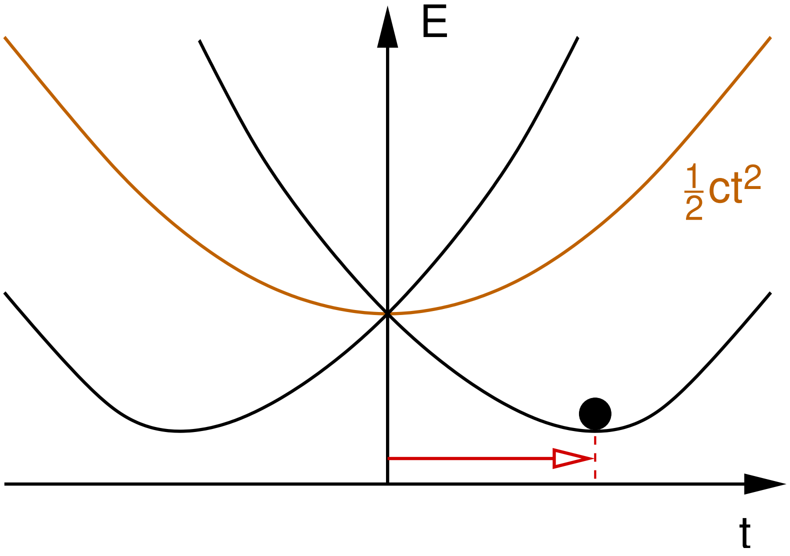


Fig. 8.1: Sketch of the energies in the Jahn-Teller effect.

8.1 Jahn-Teller Instability

The **Jahn-Teller theorem**[16] was published in 1937 and says that "any non-linear molecular system in a degenerate electronic state will be unstable and will undergo distortion to form a system of lower symmetry and lower energy thereby removing the degeneracy"

Sometimes, the Jahn-Teller effect is invisible or appears only at low temperatures. This is due to a thermal averaging over the various minima. In this case one calls it a dynamical Jahn-Teller effect.

A typical example are d^8 complexes, such as the hexaaquacopper(II) complex ion, shown in Fig. 8.2. This is a molecule that would normally exist in octahedral environment. However the doublet of e_g states is degenerate and therefore half-occupied. As result two of the bonds elongate, while the remaining four bonds contract. Thus there is a tetragonal distortion on the ion.

A good example for a Jahn-Teller instability is the vacancy in silicon. Without the Jahn-Teller distortion there are three symmetric states in the gap of silicon, which are filled by two electrons, if the defect is neutral. As a result of the Jahn teller instability two of the dangling bonds form pairwise weak bonds among each other. This leads to a lowering of one state relative to the other two. If we add two more electrons we obtain a different distortion, where two, filled, states move down in energy and one of the three orbitals moves up.

see /home/ptpb/Tree/Projects/Sivacancy

8.2 Peierl's Instability

The classical problem of a Peierl's instability is that of acetylene $H_2C[CH]_nCH_2$. The π orbitals form a band that is half filled. Thus there is a degeneracy, with half occupation at $\frac{1}{2}G = \frac{\pi}{a}$. This instability can be lifted by a structural distortion, where the carbon atoms form alternating double bond. This doubles the lattice constant and cuts the reciprocal lattice constant in half. The band structure is folded in.

Another example for a Peierls distortion is that of the silicon (001) surface. The dangling bond states form a half filled band, that can be split by forming dimers between the dangling bonds. This is the so-called dimer row reconstruction with a 2x1 surface unit cell. However also the remaining bands are half filled. The degeneracy lies again at one-half of the lattice constant. A second Peierl's

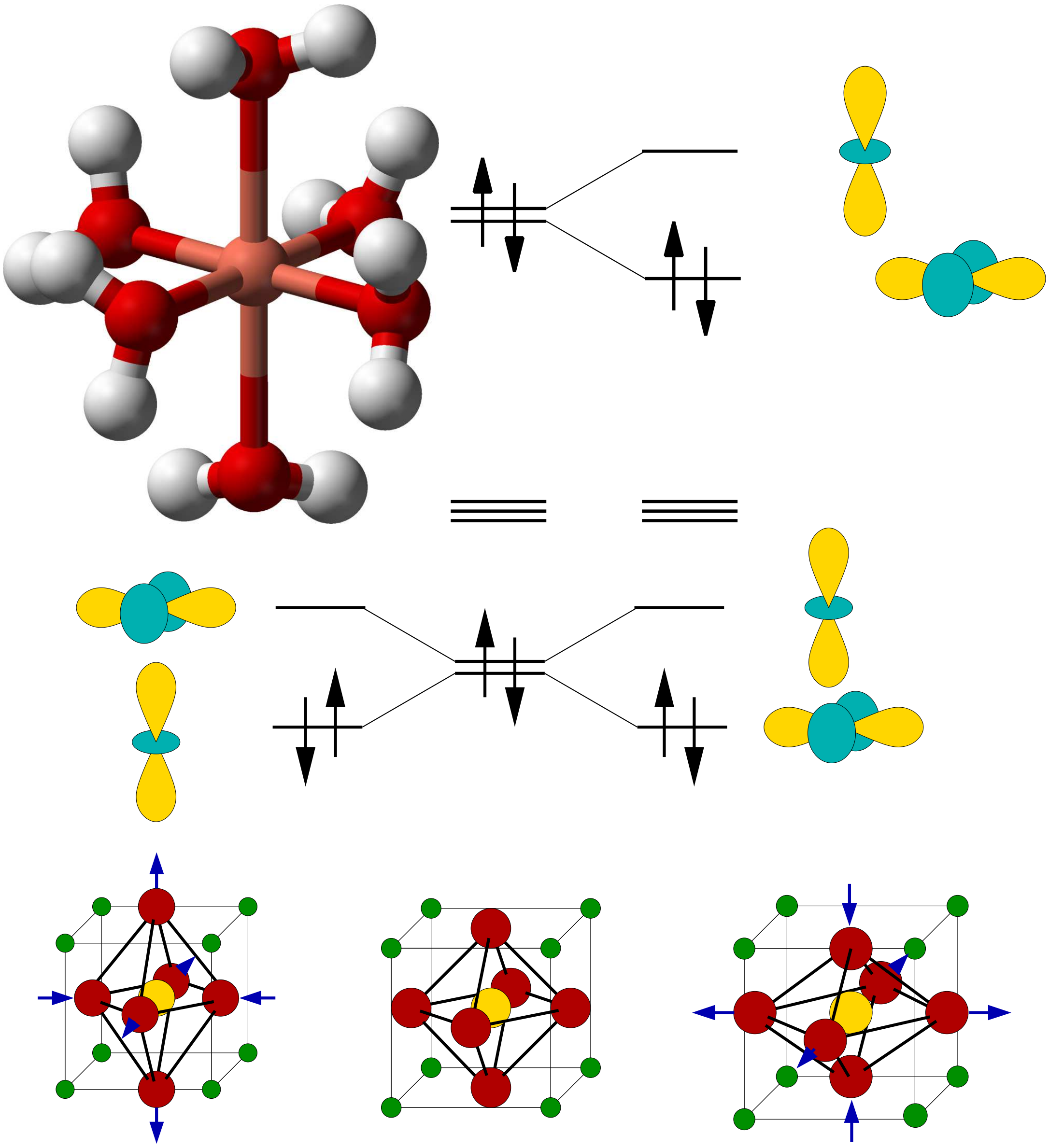


Fig. 8.2: The Jahn-Teller effect is responsible for the tetragonal distortion of the hexaaquacopper(II) complex ion, $[Cu(OH_2)_6]^{2+}$, which might otherwise possess octahedral geometry. The two axial Cu-O distances are 238 pm, whereas the four equatorial Cu¡ÝO distances are 195 pm.

distortion leads to the buckled dimer reconstruction, where the electron from one dangling bond is shifted to the other partner of the dimer bond. This occurs in an alternating manner, so that the resulting unit cell is (2x2).

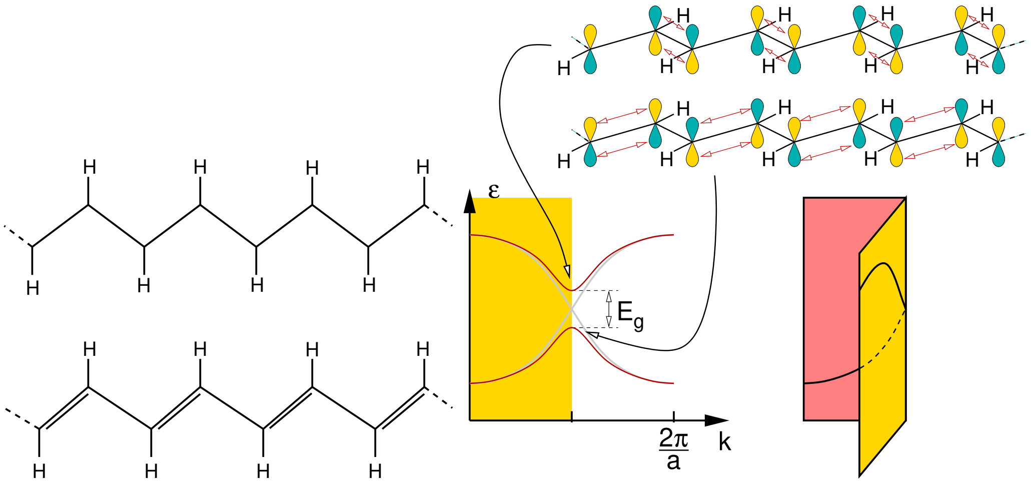


Fig. 8.3: Acetylene in the symmetric (top) and the broken symmetry state(bottom)

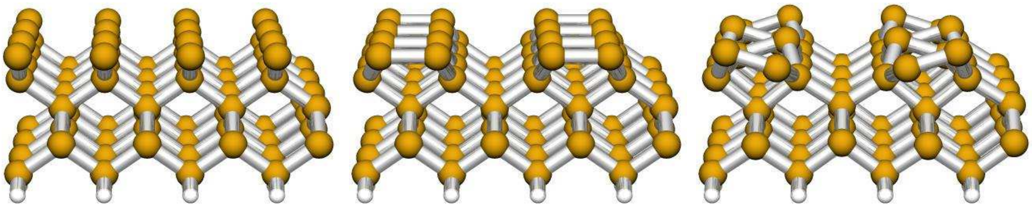


Fig. 8.4: Two subsequent Peierls distortions lead to the buckled dimer reconstruction< of the Si(001) surface

8.3 Nesting

8.4 Other instabilities

The structural instabilities are not the only ones that are possible. Instead of distorting the structure its is possible to develop a spin-polarization, that is to form a ferromagnet or an antiferromagnet, or it is possible to form a charge density wave, or even a pairing of electron pairs into Cooper pairs to form a superconductor.

Important is that a large density of states at the Fermi level is always an indication that the system would like to undergo a symmetry breaking transition. Such systems tend to exhibit interesting properties, which can be put into technological use. Which kind of transition actually occurs, however, depends a lot on the system.

Chapter 9

Ionic compounds

Besides the covalent bond the ionic bond is among the strongest. The differences between the covalent and ionic bond have been discussed by Pauling[17].

9.1 Oxidation states and charges

The concept of oxidation state has been introduced by the American chemist Wendell Latimer in 1938 in his book "The Oxidation States of the Elements and Their Potentials in Aqueous Solution". The term **oxidation state** is simply another word for charge state[1]. The term is only very indirectly related to oxygen or the process of oxidation.

A common misunderstanding is that charge states or oxidation states has something to do with the charge distribution. In fact the charge distribution is surprisingly little affected by the oxidation state. The reason is that the atomic electron of electropositive elements are located in a region where in a compound the anions would be located. If the anion is charged, its orbitals expand so that the net change in the charge distribution is minute. The charge transfer is simply screened away.[18] Thus the electron density at the cation is little affected by the anions.

The oxidation states are tremendously helpful in understanding the electron levels. Thus if one wants to understand the stability of compounds or if one wants to understand the energetic position of the energy levels the concept of formal oxidation states is extremely valuable.

9.2 Electronegativity, hardness and softness

Once the atoms are compressed, we can allow charge transfer, which is one mechanism to minimize the energy. The total energy of an atom depends not only on the radius but also its total charge.

We can expand the energy of an isolated atom in a box in a Taylor expansion in the number N of excess electrons about the neutral state. We obtain

$$E(N) = E_0 + \epsilon_0 N + \frac{1}{2} U N^2 \qquad \text{for } N_{min} \le N \le N_{max} \tag{9.1}$$

Here ϵ is the highest occupied atomic orbital and U is the Coulomb interaction parameter of that orbital. This equation is only valid as long as the same multiplet of energy levels is occupied[2]. For example, for an oxygen atom the number ΔN of excess electrons may vary from -4 to 2.

---

[1] For a definition see `http://goldbook.iupac.org/O04365.html`.

[2] In reality the total energy consists of straight line segments between integer occupations. The above form is an smooth interpolation, which may mimic broadening of the levels in an environment. Current density functionals produce a similar as in Eq. 9.1

The Coulomb interaction parameter U is also identified with **hardness**[19] η. The hardness of neutral atoms lies in the range between 2 eV and 7 eV.[19].

The connection of the energy derivative with respect to charge and the energy of the highest occupied electron level is not immediately apparent. It follows from **Janak's theorem**[20] within Density functional theory, which says

$$\epsilon_n = \frac{\partial E}{\partial f_n}$$

where ϵ_n is an one-particle level as obtained from density functional theory, E is the density functional total energy and f_n is the occupation of the level ϵ_n.

Let us consider a small model system of two atoms, that each have an energy of the form given in Eq. 9.1. In addition we consider the Coulomb interaction between the two ions in case of charge transfer. To simplify the model, we assume that the densities of the two atoms do not overlap, so that we can calculation the Coulomb interaction from a point charge model.

At sufficient large separation d, the total energy of the two atoms A and B is

$$E_{tot}(N_A, N_B) = \underbrace{E_{A,0} + \epsilon_{A,0}N_A + \frac{1}{2}U_A N_A^2}_{E_A(N_A)} + \underbrace{E_{B,0} + \epsilon_{B,0}\Delta N_C + \frac{1}{2}U_B N_B^2}_{E_B(N_B)} + \frac{e^2 N_A N_B}{4\pi\epsilon_0 d} \tag{9.2}$$

ϵ_0 is a dielectric constant. While this minimum model is extremely crude, we will explore the chemical trends that can be deduced from it.

Let us determine charge transfer and binding energy, while requiring overall charge neutrality, that is $N_A + N_B = 0$. The equilibrium condition,

$$\underbrace{\frac{\partial E_{tot}}{\partial N_A}}_{\mu_A(N_A)} + \underbrace{\frac{\partial E_{tot}}{\partial N_C}}_{\mu_B(N_B)} \underbrace{\frac{dN_B}{dN_A}}_{-1} = 0$$

implies that the chemical potentials μ_A and μ_B of the two atoms are identical. The **chemical potential** is defined as derivative of the total energy with respect to particle number

$$\mu = \frac{\partial E(N)}{\partial N}$$

The minimum total energy [3] is

---

[3] Thus we obtain the minimum of the total energy Eq. 9.2 as

$$\mu_A - \mu_B = \epsilon_A - \epsilon_B + \left[U_A + U_B - \frac{e^2}{4\pi\epsilon_0 d}\right] N_A = 0$$

$$N_A = \frac{\epsilon_B - \epsilon_A}{U_A + U_B - \frac{e^2}{4\pi\epsilon_0 d}}$$

$$\Delta E = -\frac{1}{2}\frac{(\epsilon_B - \epsilon_A)^2}{U_A + U_B - \frac{e^2}{4\pi\epsilon_0 d}}$$

If the minimum or maximum charge transfer is reached, the binding energy is

$$\Delta E = (\epsilon_A - \epsilon_B)N_X + \frac{1}{2}(U_A + U_B)N_x^2$$

The binding energy of two ions at large distance d is

$$\Delta E=-\frac{1}{2}\frac{(\epsilon_B-\epsilon_A)^2}{U_A+U_B-\frac{e^2}{4\pi\epsilon_0 d}} \qquad \text{with} \qquad N_A=-N_B=\frac{\epsilon_B-\epsilon_A}{U_A+U_B-\frac{e^2}{4\pi\epsilon_0 d}} \tag{9.3}$$

for partial charge transfer and

$$\Delta E=(\epsilon_A-\epsilon_B)N_X+\frac{1}{2}(U_A+U_B)N_x^2 \qquad \text{with} \qquad N_A=-N_B=N_X \tag{9.4}$$

for maximum charge transfer within one multiplet. The **maximum** of the two energies ΔE is the physical one.

This model shows already that the charge transfer is driven by the difference of the energies of the highest atomic orbitals. This is obvious because the system can lower its energy by placing some of the electrons into a lower level.

9.2.1 The concept of electronegativity

The concept of **electronegativity** is directly related to the position of the highest occupied electron level. Electronegativity is considered as the tendency of an atom or a molecule to attract electrons. There are different definitions of electronegativity:

Paulings electronegativities

The original definition goes back to Linus Pauling[21], who determined the electronegativity difference of two species to

$$\chi_A-\chi_B=\pm\sqrt{\left(E_d(AB)-\frac{E_d(AA)+E_d(BB)}{2}\right)\frac{1}{eV}}$$

where E_d are the bond dissociation energies. The **Pauling electronegativities** cover the range from 0.7 to 4.0.

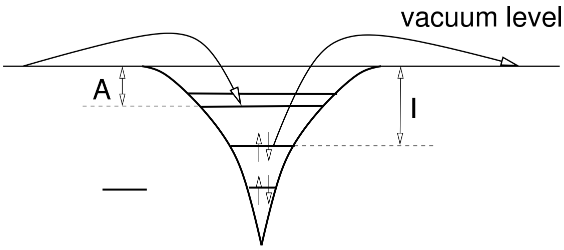
Using our simple model, we would obtain

$$\chi_A-\chi_B\approx\sqrt{\frac{1}{2}\frac{(\epsilon_B-\epsilon_A)^2}{U_A+U_D}\frac{1}{eV}}$$

Mulliken's electronegativities

Mulliken introduced the concept of "absolute electronegativity"[15] related to the mean value of electron affinity A and ionization potential I. Mulliken scaled the result in order to minimize the deviation to Pauling's electronegativity, and this posed the definition of the **absolute** or **Mulliken electronegativity** to

$$\chi=\frac{0.187}{eV}\frac{I+A}{2}$$



The **ionization potential** I is the energy required to remove an electron from a given atom or electron. Thus the ionization potential is

$$I = E[N-1] - E[N]$$

where N is the number of electrons. If N is the number of electrons for the neutral system, I is called the **first ionization potential**. If N is the number of electrons in the system with one positive charge it is called the second ionization potential. For a solid, the first ionization potential is called the **work function**. If we remove the electron , it is placed far away from the atom. The potential there is the so-called **vacuum level**, which is assumed to be the zero of the energy scale.[4]

The **electron affinity** is the energy gained, when an atom or molecule captures an additional electron.

$$A = E[N] - E[N+1]$$

Let us now return to our simple model for an atom Eq. 9.1 and determine the electro-negativity. We obtain

$$\begin{aligned} I &= -\epsilon_0 + \frac{1}{2}U \\ A &= -\epsilon_0 - \frac{1}{2}U \\ \frac{I+A}{2} &= -\epsilon_0 = -\frac{dE}{dN} \end{aligned}$$

Thus we see that the electronegativity is directly related to the position of the highest occupied orbital.

9.2.2 The concept of hardness

The second parameter in our model which determines the charge transfer in out model is the parameter U, which is what chemists call **hardness** η. Hardness is defined as

$$\eta = I - A$$

Note that the definition of hardness often carries an additional factor $\frac{1}{2}$.

Related to hardness is the **softness** which is the inverse of hardness.

If we determine the hardness in our model we obtain

$$\eta = U = \frac{d^2E}{dN}$$

---

[4] Note that the vacuum level is not a well defined quantity, if the system is infinite. The dipole of a material can cause the vacuum level to assume different values over different surfaces of an extended material.

9.2.3 Hard-Soft Acid-Base (HSAB) principle

In 1963 Pearson introduced the **HSAB principle**[22].

HSAB PRINCIPLE

All other factors being equal (namely the strength of acids and bases), hard acids prefer to bind to hard bases and soft acids prefer to bind to soft bases.

A theoretical basis for the HSAB principle has been given by Parr et al.[23].

The principle is based on the following exchange reaction between a hard acid A_h, a hard base B_h, a soft acid A_s and a soft base B_s.

$$A_hB_s + A_sB_h \Rightarrow A_hB_h + A_sB_s$$

It says that if the electronegativities of acids are equal and that of the bases are equal, the following relation holds.

$$\Delta E_{A_hB_h} + \Delta E_{A_sB_s} - \Delta E_{A_hB_s} + \Delta E_{A_sB_h} < 0$$

Following Ayers et al.[24], we may use our model Eq. 9.3 for the binding energy. The HSAB principle says that, if the electronegativities are equal, i.e.

$$\epsilon_{A_h} = \epsilon_{A_s} \qquad \text{and} \qquad \epsilon_{B_h} = \epsilon_{B_s} \, ,$$

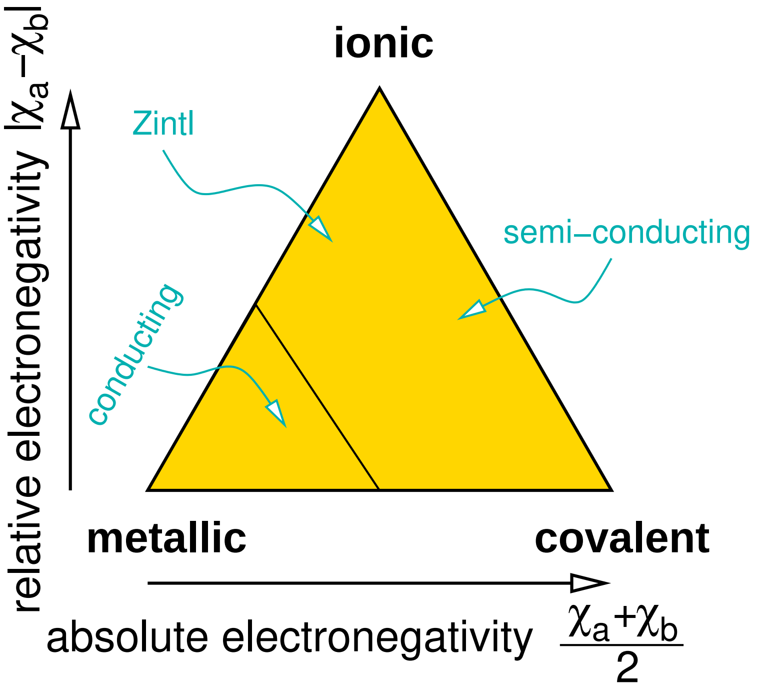
the following identity holds

$$\begin{aligned}
0 &> \Delta E_{A_hB_h} + \Delta E_{A_sB_s} - \Delta E_{A_hB_s} + \Delta E_{A_sB_h} \\
&= -\frac{1}{2}\frac{(\epsilon_{A_h}-\epsilon_{B_h})^2}{U_{A_h}+U_{B_h}} - \frac{1}{2}\frac{(\epsilon_{A_s}-\epsilon_{B_s})^2}{U_{A_s}+U_{B_s}} + \frac{1}{2}\frac{(\epsilon_{A_s}-\epsilon_{B_h})^2}{U_{A_s}+U_{B_h}} + \frac{1}{2}\frac{(\epsilon_{A_h}-\epsilon_{B_s})^2}{U_{A_h}+U_{B_s}} \\
&= -\frac{(\epsilon_A-\epsilon_B)^2}{2}\biggl(+\frac{1}{U_{A_h}+U_{B_h}} + \frac{1}{U_{A_s}+U_{B_s}} - \frac{1}{U_{A_s}+U_{B_h}} - \frac{1}{U_{A_h}+U_{B_s}}) \\
&= -\frac{(\epsilon_A-\epsilon_B)^2}{2}\biggl(\frac{U_{A_s}+U_{B_s}+U_{A_h}+U_{B_h}}{(U_{A_h}+U_{B_h})(U_{A_s}+U_{B_s})} - \frac{U_{A_s}+U_{B_h}+U_{A_h}+U_{B_s}}{(U_{A_s}+U_{B_h})(U_{A_h}+U_{B_s})}\biggr) \\
&= -\frac{(\epsilon_A-\epsilon_B)^2}{2}(U_{A_s}+U_{B_s}+U_{A_h}+U_{B_h})\biggl(\frac{1}{(U_{A_h}+U_{B_h})(U_{A_s}+U_{B_s})} - \frac{1}{(U_{A_s}+U_{B_h})(U_{A_h}+U_{B_s})}\biggr) \\
&= -\frac{(\epsilon_A-\epsilon_B)^2}{2}(U_{A_s}+U_{B_s}+U_{A_h}+U_{B_h})\biggl(\frac{(U_{A_s}+U_{B_h})(U_{A_h}+U_{B_s}) - (U_{A_h}+U_{B_h})(U_{A_s}+U_{B_s})}{(U_{A_h}+U_{B_h})(U_{A_s}+U_{B_s})(U_{A_s}+U_{B_h})(U_{A_h}+U_{B_s})}\biggr) \\
&= -\frac{(\epsilon_A-\epsilon_B)^2}{2}\frac{U_{A_s}+U_{B_s}+U_{A_h}+U_{B_h}}{(U_{A_h}+U_{B_h})(U_{A_s}+U_{B_s})(U_{A_s}+U_{B_h})(U_{A_h}+U_{B_s})} \\
&\quad\times\biggl([U_{A_s}U_{A_h}+U_{A_s}U_{B_s}+U_{B_h}U_{A_h}+U_{B_h}U_{B_s}] - [U_{A_h}U_{A_s}+U_{A_h}U_{B_s}+U_{B_h}U_{A_s}+U_{B_h}U_{B_s}]\biggr) \\
&= -\frac{(\epsilon_A-\epsilon_B)^2}{2}\frac{U_{A_s}+U_{B_s}+U_{A_h}+U_{B_h}}{(U_{A_h}+U_{B_h})(U_{A_s}+U_{B_s})(U_{A_s}+U_{B_h})(U_{A_h}+U_{B_s})} \\
&\quad\times\biggl(+U_{A_s}U_{B_s}+U_{B_h}U_{A_h}-U_{A_h}U_{B_s}-U_{B_h}U_{A_s}\biggr) \\
&= -\frac{(\epsilon_A-\epsilon_B)^2}{2}\frac{U_{A_s}+U_{B_s}+U_{A_h}+U_{B_h}}{(U_{A_h}+U_{B_h})(U_{A_s}+U_{B_s})(U_{A_s}+U_{B_h})(U_{A_h}+U_{B_s})}\underbrace{\biggl((U_{A_h}-U_{A_s})(U_{B_h}-U_{B_s})\biggr)}_{>0} \qquad \textit{q.e.d.}
\end{aligned}$$

9.3 Van Arkel Ketelaar triangle

The **Van Arkel-Ketelaar triangle** allows to classify solids according to the electronegativities of their constituents. There are two main axes:

- the difference in electronegativities is a measure of the charge transfer in a compound.
- the average electronegativity is a measure of how strongly the electrons are bound to the atomic cores.


ionic
Zintl
semi–conducting
conducting
metallic
covalent
relative electronegativity $|\chi_a - \chi_b|$
absolute electronegativity $\frac{\chi_a+\chi_b}{2}$

Let us look at the elements: Large electronegativities are found on the right-side of the periodic table. The effective core charge is large and the valence electrons are tightly bound. As a consequence, the highest electron level is relatively low. Those atoms form molecules or covalent solids. If the electronegativities are very large, there are few electrons available for bonding so that molecules are formed. If the electronegativity is somewhat smaller the materials form extended solids. Relatively large electronegativities are also found in the middle of the transition metal series.

Small electronegativities are found at the left of the periodic table. The effective core charge is small and the valence electrons are only loosely bound. Those materials tend to form metals. Metals can be seen as the positively charged ion cores, which swim in a free electron gas of the remaining electrons.

Covalent and metallic character denote the two corners of the horizontal axis of the Van Arkel-Ketelaar triangle.

The upper corner of the Van-Arkel Ketelaar triangle denotes the ionic materials, which are characterized by a large electronegativity difference of its constituents.

9.4 Zintl compounds

An interesting intermediate case are the Zintl compounds which are located on the left upper edge. One of the constituents gives up its electrons easily. The other constituent receives those electrons, but its valencies are not yet completely saturated. Thus the electropositive element forms covalently bonded networks. This is cast in the so-called Zintl rule:

The Zintl-Klemm concept says in simple terms that if we add an electron to an atom it behaves as if the atom would be shifted to the right in the periodic table, and if we remove electrons it behaves as if it were shifted to the left. That is if a element has a valence equal or greater than four, each additional electron transforms one hybrid orbital into a lone-pair, that does not form a covalent bond. Hence the covalent coordination is reduced by one for each electron added.

This rule simply reflects the observation motivated earlier that the coordination of main group elements is largely given by the number of valence electrons.

A good example for Zintl compounds are the Sr-silicides.

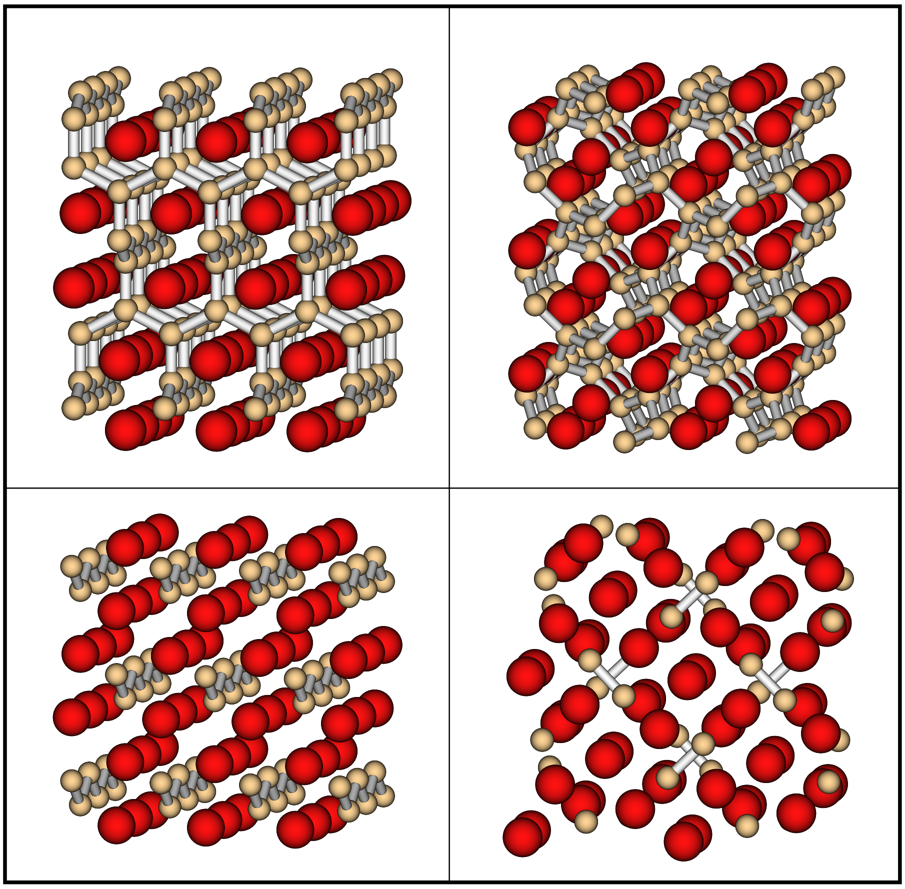


Fig. 9.1: Sr silicides: top $SrSi_2$; bottom left SrSi; bottom left Sr_5Si_3.

Silicon is a four valent element and crystallizes in the diamond structure,where each atom is tetrahedrally, that is four-fold, coordinated. If we form a silicide each strontium atom donates two electrons to the silicon network. Depending on the amount of Sr atoms, the silicon network receives a certain number of electrons. It becomes iso-electronic to phosphorus, sulfur, chlorine or argon.

- In $SrSi_2$ each silicon receives one additional electron. Hence it behaves similar to a phosphorus atom, that likes to form the bonds. Indeed the silicon network of $SrSi_2$ is a three fold coordinated network.
- In SrSi. each electron receives two electrons, and behaves like sulfur. That is it forms chains of two-fold coordinated silicon atoms.
- In Sr_5Si_3, ten electrons are distributed over three silicon atoms. This is sufficient to form two Si^{3-} ions and one Si^{4-} ion. The Si^{3-} ions have one half occupied orbital left and combine to form dimers, while the remaining Si^{4-} is isoelectronic to argon and forms an isolated ion.

The Zintl-Klemm concept is fairly general. Most of the concepts of covalent bonding where based on the number of electrons available, and on the special role of d-electrons, (f-electrons play again a special role.) When we move vertically in the periodic table the electron count remains unchanged if we move horizontally, while keeping the number of electrons unchanged, the main difference is that we add a charge to the nucleus. This charge shifts the electron levels up and down, but does not affect the nature of these electron levels. The effect is a change in the electronegativity, that is the ability of hold on to its own electrons, respectively to attract further electrons.

We observe a similar effect in molecules. For example the ammonium ion NH_4^+, the borohydride ion[5] BH_4^- are iso-electronic to methane CH_4 and hence form a tetrahedral structure.

The carbocation CH^{+3} is iso-electronic to BH_3 and forms the same structure, namely a trigonal planar structure[6]. On the other hand H_3C^- ion and ammonium NH_3 form an umbrella like a tetrahedron with a missing corner.

9.5 Madelung constants

The electrostatic energy can be approximated by a point charge model. The evaluation is non-trivial, because the lattice sums over the pairwise interactions converges poorly due to the long-ranged nature of the Coulomb interaction.

Ewald[25] proposed the **Ewald sum** technique, which divides the electrostatic energy into two parts: One is the electrostatic interaction of atom-centered spherical densities of Gaussian shape and is evaluated by a sum in reciprocal space. The other is the difference of the electrostatic interaction between point charges and these Gaussians. This results in a short ranged pair potential that can be determined by a real space sum.

For binary compounds, the electrostatic energy per ion can be expressed by a **Madelung constant** A

$$E_{mad} = \frac{1}{2\sum_i} \sum_{i,j} \frac{Q_i Q_j}{4\pi\epsilon|\vec{R}_i - \vec{R}_j|} = \frac{A}{n} \frac{Q_1 Q_2}{4\pi\epsilon_0 d} \tag{9.5}$$

where n is the number of ions in the unit cell, A is the Madelung constant Q_1 and Q_2 are the charges for cations and anions and d is the smallest distance between cations and anions. Clearly this equation is restricted to binary systems.

Unfortunately, the definition of Madelung constants is not unique and often, numbers are given without providing the underlying definition. Tables of madelung constants alongside their definitions are available[26](See supplementary material therein).

The Madelung constants for binary ionic solids is given in table 9.1. We see that the Madelung constants per ion are fairly similar. Even for a diatomic molecule made out of two dimers the value of A/n is $\frac{1}{2}$, not too far form the other values.

For NaCl, with a nearest neighbor distance of 2.814 Åwe obtain an electrostatic energy of 3.6 eV.

9.6 Born-Mayer Equation

If one includes the Pauli- or **Born repulsion** between the ions, one obtains an estimate for the complete energy of a ionic crystal.[28] One can use an exponential ansatz for the Pauli repulsion, namely

$$E_{rep} = b\mathrm{e}^{-\frac{d}{\rho}}$$

where $\rho = 0.345$ Å is a approximate constant determined empirically.

The constant b can be determined from the lattice constant. The total energy is

$$E(d) = b\mathrm{e}^{-\frac{d}{\rho}} + \frac{A}{n} \frac{Q_1 Q_2}{4\pi\epsilon_0 d} \tag{9.6}$$

---

[5] see sodium borohydrate $NaBH_4$

[6] borane dimerizes to form B_2H_6.and exists in its molecular form only in the gas phase.

	A	n	$\frac{A}{n}$
sodium chloride	1.74756	2	0.8878
cesium chloride	1.76267	2	0.881335
fluorite	2.5194	3	0.8398
corundum	4.1719	5	0.83438
wurtzite	1.64132	2	0.82065
zincblende	1.63805	2	0.81925
sphalerite	1.6381	2	0.81905
rutile	2.408	3	0.802667
cuprite	2.0577	3	0.6859
anatase	2.400	3	0.8
CdI_2	2.35	3	0.78333
β-quartz	2.2197	3	0.7399
dimer	1	2	0.5

Table 9.1: Madelung constants for binary ionic solids[27]. $E = A\frac{Q_AQ_B}{4\pi\epsilon_0|d|}$ where d is the smallest anion-cation distance. (Not all values have been taken from Pauling's book)

At the equilibrium distance the force must vanish, that is

$$\left.\frac{dE}{dd}\right|_{d_0} = -\frac{b}{\rho}e^{-\frac{d_0}{\rho}} - \frac{A}{n}\frac{Q_1Q_2}{4\pi\epsilon_0 d_0^2} = 0$$

$$b = -\frac{A}{n}\frac{Q_1Q_2}{4\pi\epsilon_0 d_0^2}\rho e^{+\frac{d_0}{\rho}}$$

Now we insert b into Eq. 9.6 and obtain

$$E = \frac{A}{n}\frac{Q_1Q_2}{4\pi\epsilon_0 d_0}\left(1-\frac{\rho}{d_0}\right) \tag{9.7}$$

This is the so-called **Born-Mayer equation**.

For NaCl, the shortest distance d_0 between ions is 2.82 Å, which leads to a reduction of the total energy relative to the electrostatic energy by about 12 %.

9.7 Kapustinskii Equation

Kapustinskii[29] recognized that the Madelung constant divided by the number of ions in the unit cell leads to nearly the same value for a variety of structures. The resulting value increases with coordination number. However since also the distance increases with coordination the result balances out. That is he postulates that

$$\frac{A}{n}$$

is a constant. A is the Madelung constant, n the number of ions in the unit cell and $r_A + r_B$ is the sum of the ionic radii.

Kapustinskii estimates the minimum distance from the ionic radii and replaces the Madelung constant by that of the NaCl structure (A/n=0.8878). The result is the **Kapustinskii equation**

$$\frac{E}{n} = 0.8878\frac{Q_AQ_B}{4\pi\epsilon_0(r_A+r_B)}\left(1-\frac{\rho}{r_A+r_B}\right) \tag{9.8}$$

which allows one to estimate the lattice energy even without knowing the structure! The agreement with the lattice energy for ionic materials is typically smaller than 5 %.

9.8 Born-Haber Cycle

The Born Haber cycle allows one to estimate the energy of reaction, when an ionic solid is formed from two elemental substances. Parts of the Born-Haber cycle can also be used to estimate the cohesive energy of an ionic substance.

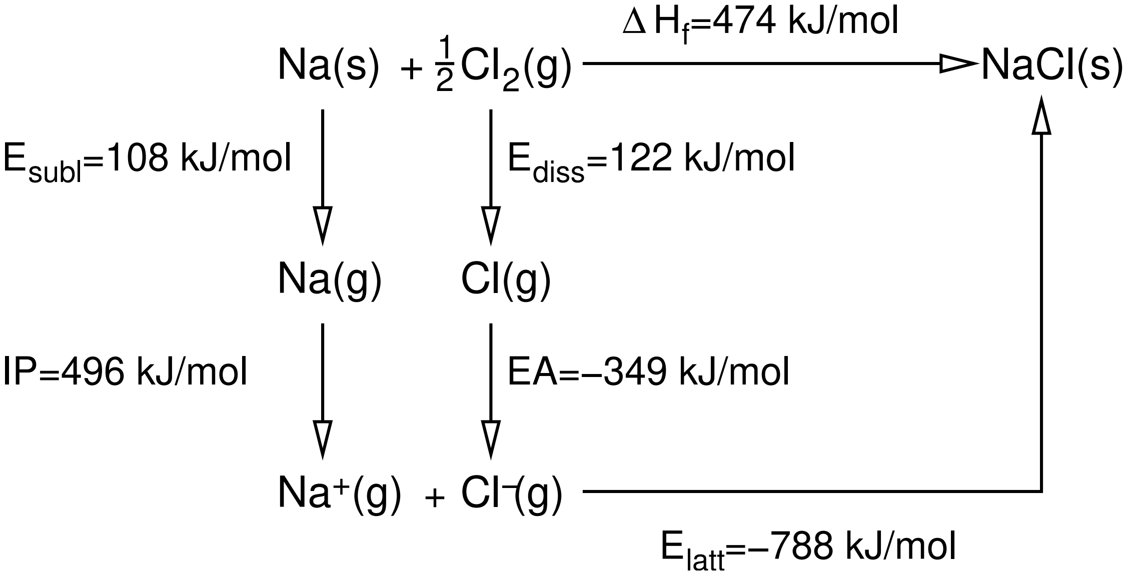


Fig. 9.2: Born-Haber cycle of NaCl.

In the following we will use the energy unit kJ/mol. The conversion is 1 eV=96.485309 kJ/mol. For simple estimates one simply uses the factor 1 eV≈ 100 kJ/mol. Part of the calculation will be done in Hartree atomic units, defined by $\hbar = e = m_e = 4\pi\epsilon_0 = 1$. The unit of energy in the Hartree atomic unit system is the Hartree: 1 H=27.211396 eV=2625.5 kJ/mol. The unit of length in the Hartree atomic unit system is the Bohr radius 1 a_0=0.5291772 Å=0.5291772 $\times 10^{-10}$ m.

Let us consider a simple example:

1. We start out from metallic sodium and gaseous chlorine. First we need to decompose the elemental materials into its atoms. In the case of chlorine this is **heat of dissociation**, which is 122 kJ/mole for Cl_2. For the metal it is the **sublimation energy** 108 kJ/mol. Per formula unit we need to supply $\frac{1}{2} \cdot 122 + 108$ kJ/mol=169 kJ/mol, in order to form isolated atoms.

2. The next step is to convert the atoms into the respective ions. The energy to remove an electron from an atom and place it at infinity is the first ionization potential. For Na it has the value of 496 kJ/mol. The energy to add an electron to an atom is the electron affinity. For Cl it has the value of -349 kJ/mol. Thus we need to supply 147 kJ/mol to convert the atoms into ions.

3. Finally we need to bring the ions together to form the ionic solid. The lattice energy of NaCL is $-788\ kJ/mol$.

4. If we add the three energy contributions together we obtain the heat of formation of sodium chloride, namely 474 kJ/mol.

Let us investigate the lattice energy.

- Using only the Madelung energy of sodium chloride,we obtain

  $$E_{latt} = n\frac{A}{n} \times \frac{Q_1Q_2}{4\pi\epsilon_0 d} = 0.3279H = 861kJ/mol.$$

  The madelung constant for sodium chloride structure is $A = 1.74756$. The number of ions per formula unit is $n = 2$. The interatomic distance is $d_{Na-Cl} = 2.82$ Å=5.329 a_0.In atomic units $e = 4\pi\epsilon_0 = 1$. Thus we obtain in Hartree atomic units an energy of $E_{latt} = 0.3279$ H.(H stands for Hartree, the energy unit of the Hartree atomic unit system. 1H=2625.5 kJ/mol.)

- Using the Born-Mayer equation Eq. 9.8 we obtain $E_{latt} = 756$ kJ/mol. The error is 30 kJ/mol, that is about 0.3 eV.

9.9 Pauling bonding rules

Another approach to think about coordination has been given by Pauling[7] Pauling set up a number of rules that hold pretty well for ionic compounds[32]. The rationale behind is that ions would like to pack together as much as possible, in order to maximize the electrostatic attraction of opposite ions. However, each ion has a natural radius at which the wave functions overlap so much that they repel each other appreciably. Thus, the structure according to this model is given by a balance between Pauli repulsion of the electronic shells and the Coulomb attraction on the other. The stronger the covalent contribution, the less reliable are these rules.

- **Pauling's first rule: "A coordinated polyhedron of anions is formed about each cation, the cation-anion distance being determined by the radius sum, and the ligancy of the cation by the radius ratio."** (See [27], p544.)

  The rationale is that if the anions are too large, the Pauli repulsion between the anions would prevent them to come close to the cation.

  The ratio of the radii in Tab. 9.2 is calculated assuming the ions as hard spheres. The optimum structures have touching cation-anion distances so that the electrostatic interaction is maximized. On the other hand the anions must not overlap due to their Pauli repulsion.

  - trigonal planar: The centers of the anions form an equilateral triangle which has angles of 60°.

    $$\frac{r_A}{r_C + r_A} = \cos(30^\circ) = \frac{1}{2}\sqrt{3} \qquad ; \qquad \frac{r_C}{r_A} = \frac{2}{\sqrt{3}} - 1$$

  - tetrahedral: the rectangle is a diagonal place through the cube. The base has a length of $\sqrt{2}a = 2r_A$. Cation and anion are connected by a main diagonal through the cube. Therefore $r_A + r_C = \frac{1}{2}\sqrt{3}a$.

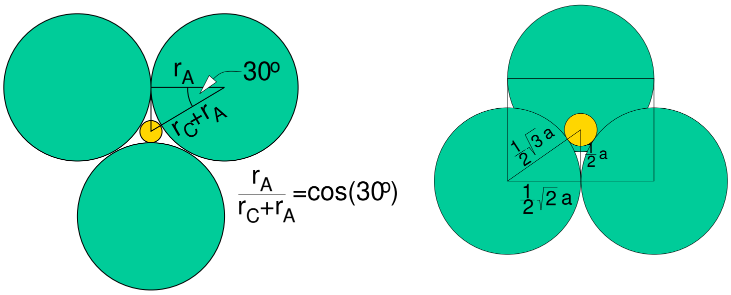
    $$\frac{r_A + r_C}{r_A} = \frac{\frac{1}{2}\sqrt{3}a}{\frac{1}{2}\sqrt{2}a} = \sqrt{\frac{3}{2}} \qquad ; \qquad \frac{r_C}{r_A} = \sqrt{\frac{3}{2}} - 1$$

  - octahedral or square planar:

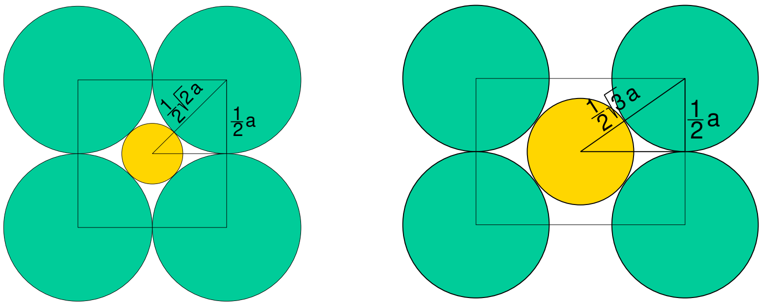
    $$\frac{r_A + r_C}{r_A} = \frac{\frac{1}{2}\sqrt{2}a}{\frac{1}{2}a} = \sqrt{2} \qquad ; \qquad \frac{r_C}{r_A} = \sqrt{2} - 1$$

---

[7] Apparently[30], the radius ratio's, which are part of the rules go back to the Austrian chemist G.F. Hüttig[31].



trigonal planar tetrahedral



octahedral or square planar cube

$0 < \frac{r_c}{r_a}$	2	linear
$0.155 < \frac{r_c}{r_a}$	3	triangular
$0.255 < \frac{r_c}{r_a}$	4	tetrahedral
$0.414 < \frac{r_c}{r_a}$	6	octahedral
$0.414 < \frac{r_c}{r_a}$	6	square planar
$0.645 < \frac{r_c}{r_a}$	8	square anti-prism
$0.732 < \frac{r_c}{r_a}$	8	cubic
$0.732 < \frac{r_c}{r_a}$	9	
$1.0 < \frac{r_c}{r_a}$	12	cubo-octahedral

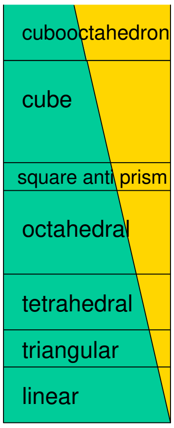


Table 9.2: Structure types as function of the ratio between the cation ionic radius r_c and the anion ionic radius r_a[27, 33]. The square anti-prism is a cube with the top rotated about 45° about the vertical axis. (See [27], p545.) At the bottom right the ranges of various structure types is shown for the ratio of anion- (green) and cation-radii (yellow).

  - square anti prism. The square antiprism is like a cube with the top face rotated by 45° about the vertical axis. Because the anions move closer together the space for the cation is smaller than in the cube.
  - cube: Shown is the diagonal plane through the cube.

$$\frac{r_A + r_C}{r_A} = \frac{\frac{1}{2}\sqrt{3}a}{\frac{1}{2}a} = \sqrt{3} \qquad ; \qquad \frac{r_C}{r_A} = \sqrt{3} - 1$$

- **Pauling's second rule: In a stable ionic structure, the valence of each anion, with changed sign is exactly or nearly equal to the sum of the strengths of the electrostatic bonds to it from the adjacent cations.**(See [27], p548.)

  In other words: For each polyhedron, let us attribute to each corner an equal fraction of the positive charge of the cation in its center. If several polyhedra share the same corner, the charges from the contributing polyhedra are summed up. According to Pauling's second rule, the resulting charge is oppositely equal to that of the anion at this corner.

Pauling calls this rule the "electrostatic valence rule". This rule corresponds to the requirement of local charge neutrality.

- **Pauling's third rule: The presence of shared edges and especially shared faces in a coordinated structure decreases its stability; this effect is large for cations with large valence and small ligancy."** (See [27], p559.) In other words, Corner-joined polyhedra are more favorable than edge sharing polyhedra, which, in turn, are more favorable than face-sharing polyhedra.

  This rule is justified because the ions of same character, say the cations in the center of the polyhedra, would like to stay as far away from each other as possible in order to minimize their electrostatic repulsion.

- **Pauling's fourth rule: "In a crystal containing different cations, those with large valence and small coordination number tend not to share polyhedron elements with each other"** (See [27], p561.)

- **Pauling's fifth rule: Structures will form from only few component elements.**

  The rationale behind this weakest rule of Pauling is that for a material with many components it is likely that there is a phase separation, which withdraws one element from the mixture.

  I did not find this rule in the Book by Pauling. It is mentioned however in [33]

Chapter 10

Bond strengths

10.1 Covalent bonds

Single bonds have bond-energies in the range of 2-4 eV. (200-400 kJ/mol).

H-H	436 kJ/mol	4.52 eV	CCsingle	414 kJ/mol	4.29 eV
C-H	414 kJ/mol	4.29 eV	CCdouble	263 kJ/mol	2.73 eV
N-H	389 kJ/mol	4.12 eV	CCtriple	226 kJ/mol	2.34 eV
O-H	464 kJ/mol	4.81 eV	NN single	163 kJ/mol	1.69 eV
F-H	568 kJ/mol	5.89 eV	NN double	255 kJ/mol	2.64 eV
F-H	568 kJ/mol	5.89 eV	NN triple	527 kJ/mol	5.46 eV
Cl-H	431 kJ/mol	4.47 eV	CNsingle	305 kJ/mol	3.16 eV
Br-H	365 kJ/mol	3.78 eV	CNdouble	310 kJ/mol	3.21 eV
I-H	299 kJ/mol	3.09 eV	CNtriple	275 kJ/mol	2.85 eV

Table 10.1: Bond energies ΔH at 25°. From[34]. The energies for double bond are the energies between a a double bonded system and a comparable single bonded system.

10.2 Ionic bonds

The electrostatic contribution to a bond may be estimated from the Kapustinskii Equation Eq. 9.8.

For an dimer having one positive and one negative charge we obtain an energy of about 2. eV assuming a bond length of 3 Åand the constant $A/n = \frac{1}{2}$.

For an ionic crystal such as NaCl we obtain an energy per ion of about 4 eV. However, we need to divide this result by the number of bonds per atom. If we call each nearest ion-ion distance a bond, there are three bonds per ion resulting in a bond energy of 1.4 eV.

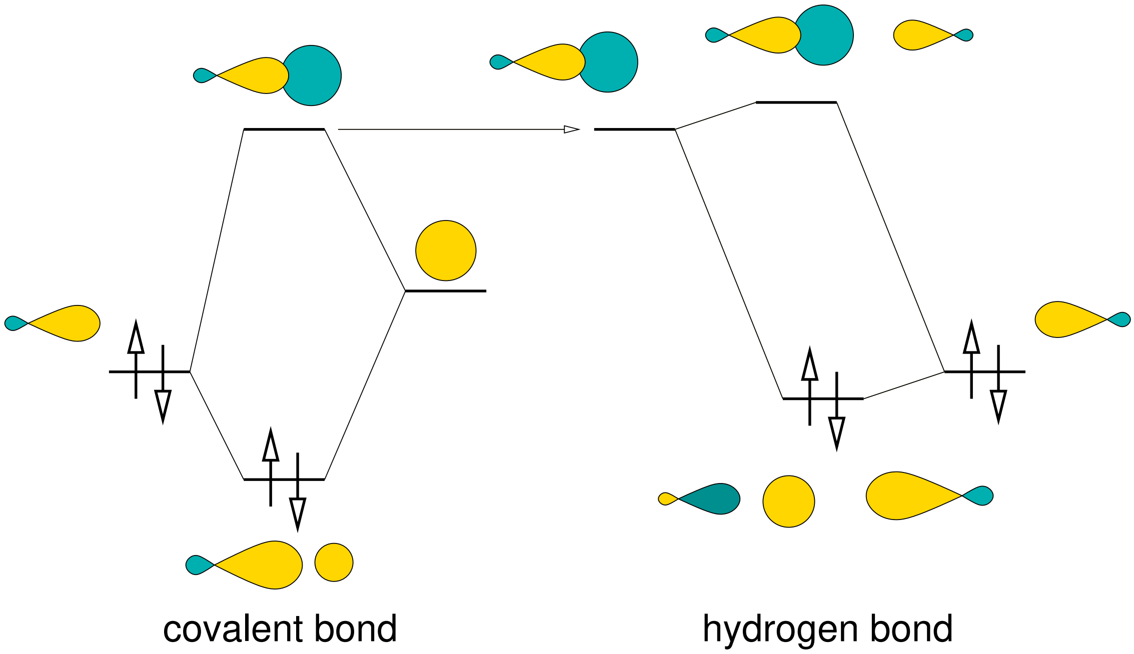
Chapter 11

Weak Bonds

11.1 Hydrogen bonds

The hydrogen bond is what keeps water molecules together in a liquid or in ice. Hydrogen bonds are essential to give proteins their shape.

The hydrogen bond is a weak bond between a lone-pair and the anti-bonding orbital of a main group element and a hydrogen atom. The hydrogen bond is special because hydrogen exhibits a non-directional s-orbital.



The hydrogen bond in water (ice) has a bond enthalpy of 25 kJ/mol[35]. Normally the hydrogen bond strength vary between 10 and 50 kJ/mol. An exceptionally strong hydrogen bond is that in fluoric acid HF, which has an enthalpy of 161 kJ/mol.[35]

11.2 Van der Waals Bond

The **van-der-Waal's bond**is the weakest bond, but it is also abundant. Compared to an ionic bond, or a covalent bond it is clearly negligible. However, in proteins and polymers, the covalently bonded network is tied together by hydrogen bonds, electrostatic attraction and van der Waals bonds.

Van der Waal's interaction has a $\frac{1}{|\vec{R}-\vec{R}'|^6}$ distance dependence, in the limit of long distances. At short distances the Pauli repulsion kicks in and forms a lower cutoff. The distance where van der Waal's attraction and Pauli repulsion are in equilibrium is the **van der Waal's distance**.

The van der Waal's bond is due to electrostatic interaction between the quantum mechanical fluctuations of the electrostatic dipoles on two polarizable molecules. While the average electrostatic dipoles vanish, the quantum fluctuations become synchronized, so that a net attractive interaction results.

Let us first make s simple model of a polarizable molecule. We treat the molecule on the most simple level, namely a point dipole. Consider a molecule in an electric field. The atom will be polarized, that is it produced a dipole that can interact with the electric field. The molecule is polarizable. The energy of a polarizable molecule in an electric field can be written[1] as

$$E(\vec{p}) = \frac{1}{2\alpha}\vec{p}^2 - \vec{E}\vec{p}$$

where α is the electric polarizability of the molecule. The ground state is given by $\frac{dE}{dp_i} = 0$ that is by

$$\vec{p} = \alpha\vec{E}$$

Now let us work out the electrostatic interaction between two electrostatic point dipoles. The electrostatic potential of a point dipole is

$$\Phi(\vec{r}) = \frac{\vec{r}\vec{p}}{4\pi\epsilon_0|\vec{r}|^3}$$

The interaction of a second dipole $\vec{p}'$at position $\vec{R}$ is

$$\begin{aligned} E_{int}(\vec{r}) &= -\vec{p}'\vec{E}(\vec{R} = -\vec{p}'\ \vec{\nabla}\big|_{\vec{R}} \frac{\vec{r}\vec{p}}{4\pi\epsilon_0|\vec{r}|^3} = -\frac{\vec{p}'\vec{p}}{4\pi\epsilon_0|\vec{R}|^3} + 3\frac{(\vec{r}\vec{p})(\vec{r}\vec{p}')}{4\pi\epsilon_0|\vec{R}|^5} \\ &= \frac{3(\vec{R}\vec{p})(\vec{R}\vec{p}') - (\vec{p}'\vec{p})\vec{R}^2}{4\pi\epsilon_0|\vec{R}|^5} \end{aligned}$$

Thus we obtain the energy of two polarizable molecules, that can electrostatically interact with each other as

$$E(\vec{p}, \vec{p}', \vec{R}) = \frac{\vec{p}^2}{2\alpha} + \frac{\vec{p}'^2}{2\alpha'} + \frac{3(\vec{R}\vec{p})(\vec{R}\vec{p}') - (\vec{p}'\vec{p})\vec{R}^2}{4\pi\epsilon_0|\vec{R}|^5}$$

We see immediately that none of the molecules is polarized and there is no interaction between the molecules.

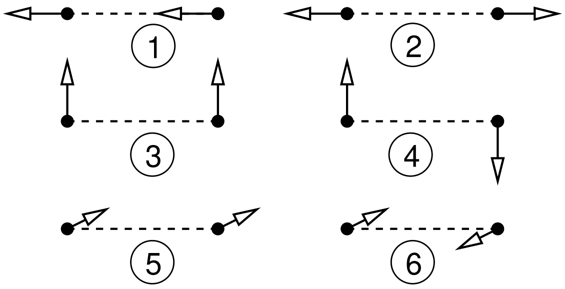
However we need to quantize the harmonic oscillator. The ground state energy of a quantum mechanical harmonic oscillator is given as $\frac{1}{2}\hbar\omega$, which is just the zero-point energy. The zero point fluctuations will now lead to a fluctuating electric field, on the other molecule. The fluctuations of the two molecules become synchronized, leading to an effective interaction.

---

[1]

$$\begin{aligned} \rho_{dip}(\vec{r}) &= -\vec{p}\vec{\nabla}\delta(\vec{r}-\vec{r}_0) \\ E_{int} &= \int d^3r\ \rho_{dip}(\vec{r}) \underbrace{\left[-\vec{E}\vec{r}\right]}_{\Phi(\vec{r})}] = \int d^3r\ \left[-\vec{p}\vec{\nabla}\delta(\vec{r}-\vec{r}_0)\right] \left[-\vec{E}\vec{r}\right] \\ &\overset{Gau}{=} \int d^3r\ \left[\delta(\vec{r}-\vec{r}_0)\vec{p}\vec{\nabla}\right] \left[-\vec{E}\vec{r}\right] = -\int d^3r\ \delta(\vec{r}-\vec{r}_0)\vec{p}\left[\vec{\nabla}\otimes\vec{r}\right]\vec{E} \\ &= -\int d^3r\ \delta(\vec{r}-\vec{r}_0)\vec{p}\vec{E} = -\vec{p}\vec{E} \end{aligned}$$

We need to derive the frequencies of the system of two dipoles. This can be done classically. Let us decompose the energy into its eigenmodes. We can use symmetry to decompose the eigenvectors of the Hessian matrix $\frac{d^2E}{dp_i dp_j}$.



$$
\begin{aligned}
E_1(x) &= \frac{1}{\alpha}x^2 + \frac{2x^2}{4\pi\epsilon_0|\vec{R}|^3} &\quad \text{with} \quad & \vec{p} = x\vec{e}_x; \vec{p}' = x\vec{e}_x \\
E_2(x) &= \frac{1}{\alpha}x^2 - \frac{2x^2}{4\pi\epsilon_0|\vec{R}|^3} &\quad \text{with} \quad & \vec{p} = -x\vec{e}_x; \vec{p}' = x\vec{e}_x \\
E_3(x) &= \frac{1}{\alpha}\alpha^2 - \frac{x^2}{4\pi\epsilon_0|\vec{R}|^3} &\quad \text{with} \quad & \vec{p} = x\vec{e}_z; \vec{p}' = x\vec{e}_z \\
E_4(x) &= \frac{1}{\alpha}\alpha^2 + \frac{x^2}{4\pi\epsilon_0|\vec{R}|^3} &\quad \text{with} \quad & \vec{p} = -x\vec{e}_z; \vec{p}' = x\vec{e}_z \\
E_5(x) &= \frac{1}{\alpha}\alpha^2 - \frac{x^2}{4\pi\epsilon_0|\vec{R}|^3} &\quad \text{with} \quad & \vec{p} = x\vec{e}_y; \vec{p}' = x\vec{e}_y \\
E_6(x) &= \frac{1}{\alpha}\alpha^2 + \frac{x^2}{4\pi\epsilon_0|\vec{R}|^3} &\quad \text{with} \quad & \vec{p} = -x\vec{e}_y; \vec{p}' = x\vec{e}_y
\end{aligned}
$$

These equations give us right away the force constants for all six modes. if we assume an isotropic mass m for the dipole dynamics, we obtain the zero-point energy as

$$
\begin{aligned}
E &= \sum_{i=1}^{6} \hbar\omega_i = \sum_{i=1^6} \hbar\sqrt{\frac{c_i}{m}} \\
&= \frac{\hbar}{\sqrt{\alpha m}} \left[\sqrt{1 + \frac{2\alpha}{4\pi\epsilon_0|\vec{R}|^3}} + \sqrt{1 - \frac{2\alpha}{4\pi\epsilon_0|\vec{R}|^3}} + 2\sqrt{1 + \frac{\alpha}{4\pi\epsilon_0|\vec{R}|^3}} + 2\sqrt{1 - \frac{\alpha}{4\pi\epsilon_0|\vec{R}|^3}} \right)
\end{aligned}
$$

For long distances $|R|$ we can consider $\frac{\alpha}{4\pi\epsilon|\vec{R}|^3}$ as small quantity and restrict ourselves to the leading order term in the Taylor expansion. The Taylor expansion of the square root is $\sqrt{1+\Delta} = 1 + \frac{1}{2}\Delta - \frac{1}{8}\Delta^2 + O(\Delta^3)$. We obtain

$$
E(\vec{R}) = \frac{\hbar}{\sqrt{\alpha m}} \left[6 - \frac{3}{4} \left(\frac{\alpha}{4\pi\epsilon|\vec{R}|^3} \right)^2 + O(\frac{1}{|R|^9}) \right]
$$

The constant is irrelevant and the distance dependent term is the Van der Waals interaction

$$
\begin{aligned}
E_{VdW}(|\vec{R}|) &= -\frac{\hbar}{\sqrt{\alpha m}} \frac{3}{4} \left(\frac{\alpha}{4\pi\epsilon_0|\vec{R}|^3} \right)^2 \\
&= -\frac{3\hbar\alpha^{\frac{3}{2}}}{4(4\pi\epsilon_0)^2\sqrt{m}} \frac{1}{|\vec{R}|^6}
\end{aligned}
$$

In this formula the mass has no physical significance other that $\omega = \sqrt{\frac{c}{m}}$ is the frequency of the dipole oscillation. The dipole oscillation frequency can be related to the optical excitation frequency of the isolated molecule, respectively the band gap. Thus we find the approximate relation

$$\hbar\sqrt{\frac{c}{m}} = E_g$$

The force constant is that of the isolated molecule, that is $\frac{1}{\alpha}$, which yields

$$\sqrt{m} = \frac{\hbar}{E_g\sqrt{\alpha}}$$

This leads to the final expression for the Van der Waal's energy

$$E_{VdW}(|\vec{R}|) = -\frac{3E_g\alpha^2}{4(4\pi\epsilon_0)^2}\frac{1}{|\vec{R}|^6}$$

Chapter 12

Crystal structures

The zoo of crystal structures can be very confusing. At first sight the atoms seem to be able to organize in any possible arrangement. However there are some characteristics, that allows us to bring some order into the chaos.

As with the electronic states there is a long-ranged view and a short ranged view. The long-ranged view of crystals are the symmetry groups and lattice types. The short ranged view discusses coordination polyhedra, bond numbers etc.

Here we will first provide the general principles, and then we will discuss the crystal structures in the light of these patterns.

12.1 Online resources

The following is a small selection of sources that, I use frequently. The selection is absolutely subjective and it is not meant to be complete in any sense.

- Prototypes of crystal structures can be fund in the AFLOW library of Crystallographic prototypes[36, 37] `http://aflowlib.org/CrystalDatabase/prototype_index.html`.
- An excellent source on cystallography is the Bilbao Crystallographic Server `http://www.cryst.ehu.es`

12.2 Bravais lattices

There are 14 different lattice types in three dimensions. They are called **Bravais lattices** after Auguste Bravais[1]

In a crystal the point groups symmetry is limited by the requirement of discrete translational symmetry. For example it is not possible to form a crystal with a five-fold symmetry axis. This does not imply that it is not possible to form materials with a five fold symmetry, but they cannot be described by three translational lattice vectors. Such materials are called **quasi crystals**.

There are 14 Bravais lattices, which can be attributed to 7 crystal systems. For a given crystal system, there is always a primitive lattice, that is a lattice point at each corner of the unit cell, and variants such as face-centered, body centered and base centered lattices.

- the cubic system has four three-fold rotation axes
- the hexagonal system has one six-fold rotation axes

---

[1] Auguste Bravais, 1811-1863. French Physicist

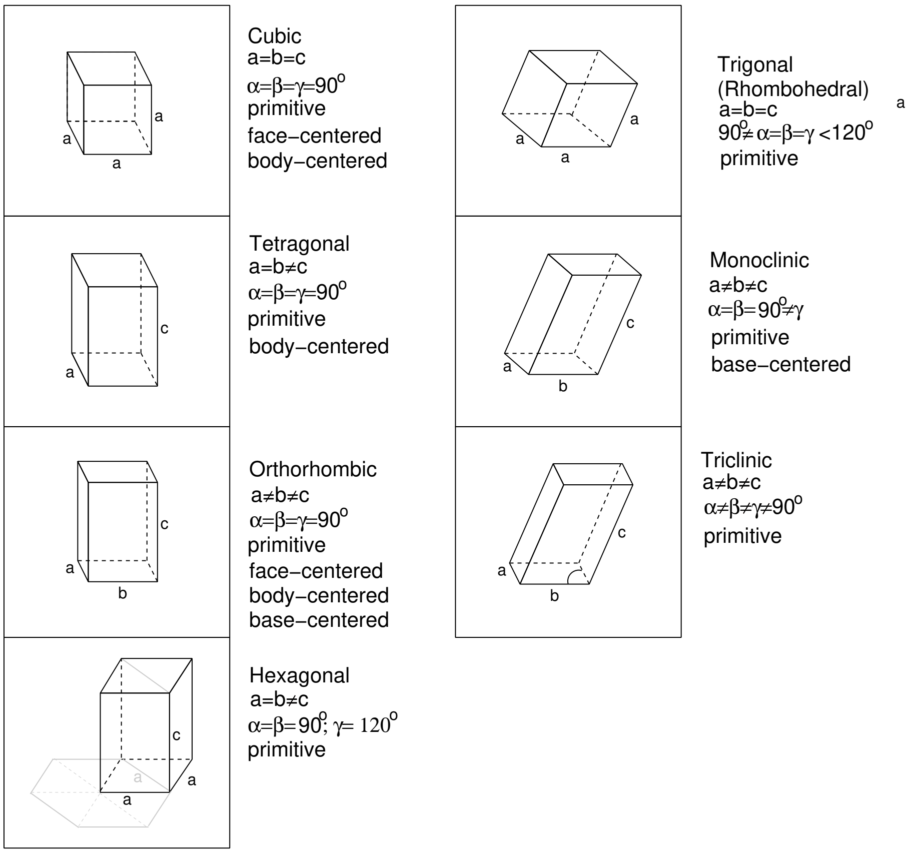

a
a
a
Cubic
a=b=c
α=β=γ=90°
primitive
face−centered
body−centered
a
a
a
Trigonal
(Rhombohedral)
a=b=c
90°≠α=β=γ<120°
primitive
a
c
a
Tetragonal
a=b≠c
α=β=γ=90°
primitive
body−centered
c
a
b
Monoclinic
a≠b≠c
α=β=90°≠γ
primitive
base−centered
c
a
b
Orthorhombic
a≠b≠c
α=β=γ=90°
primitive
face−centered
body−centered
base−centered
c
a
b
Triclinic
a≠b≠c
α≠β≠γ≠90°
primitive
c
a
a
a
Hexagonal
a=b≠c
α=β=90°; γ= 120°
primitive

Fig. 12.1: The fourteen Bravais lattices. The angle α is the one between the lattice vectors b and c, the angle β is the one between a and c. and the angle γ is the one between a and b

- the trigonal (or rhombohedral) system has one three-fold rotation axes
- the tetragonal system has one four-fold rotation axis
- the orthorhombic system requires either three twofold axes of rotation or one twofold axis of rotation and two mirror planes.
- the monoclinic system requires either one twofold axis of rotation or one mirror plane.
- the triclinic system contains no other symmetry than the translational symmetry, except the inversion.

Depending on the symmetries of the lattices, that are broken by the atoms in the unit cell, one can construct 230 distinct space groups.

12.3 Coordination polyhedra

The hybrid orbitals already showed us that certain coordinations are favorable in certain situations. The favorable coordination depends on the number of electrons and the number of bonds formed. Therefore its is useful, to look at molecules and materials is terms of local coordination of the atoms[33].

Fig. 12.2 shows the most important prototypes for a environment of an atom.

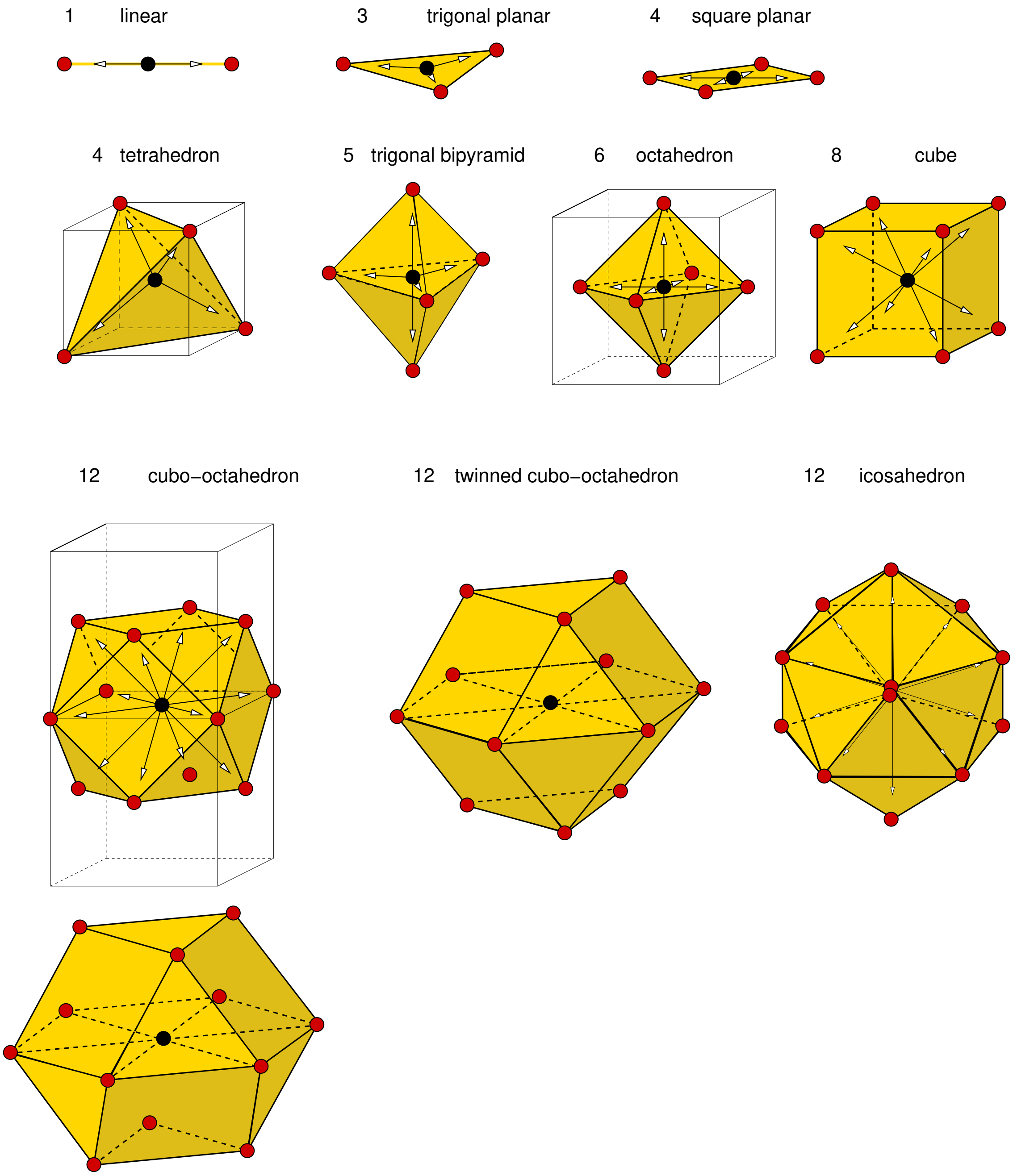

1 linear
3 trigonal planar
4 square planar
4 tetrahedron
5 trigonal bipyramid
6 octahedron
8 cube
12 cubo−octahedron
12 twinned cubo−octahedron
12 icosahedron

Fig. 12.2: Coordination polyhedra. The cubo-octahedron can be seen as an octahedron with the corners chopped off.

- tetrahedral coordination: As we will see below, a close sphere packing has octahedral and tetrahedral voids. the four sp^3 orbitals point towards the neighbors in a tetrahedral environment.
- trigonal bipyramid: five-valent elements such as phosphorus tend to form trigonal bipyramids. In the axial direction there is a three-center bond, and in the equatorial directions there are three sp^2 hybrid orbitals. Example: PH_5.
- octahedral coordination: As we will see below, a close sphere packing has octahedral and

tetrahedral voids. The two eg orbitals of the d-shell point in the direction of otahedral neignbors.

- cubo-octahedron: the cubo-octahedron is the coordination of a face-centered cubic lattice, one type of close sphere packing.

- twinned cubo-octahedron: the twinned cubo-octahedron is the coordination of a hexagonally close packed lattice, one type of close sphere packing.

12.4 Simple cubic structure (sic)

The simple cubic structure is rare in nature. One material crystallizing in the simple cubic structure is α-Polonium.

$\vec{T}_1$	(a,0,0)
$\vec{T}_2$	(0,a,0)
$\vec{T}_3$	(0,0,a)
$\vec{R}$	(0,0,0)

The simple cubic structure by itself is not very relevant in nature, but is the starting for the more common cubic structure such as the face-centered cubic (fcc) structure and the body-centered cubic (bcc) structure.

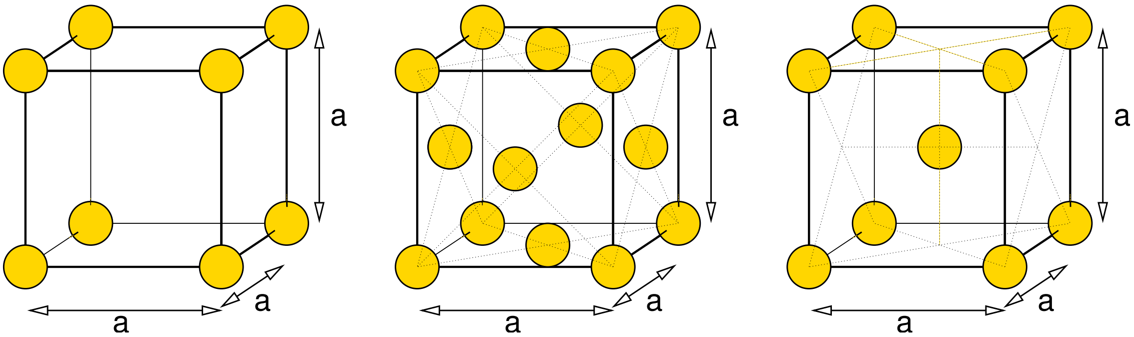
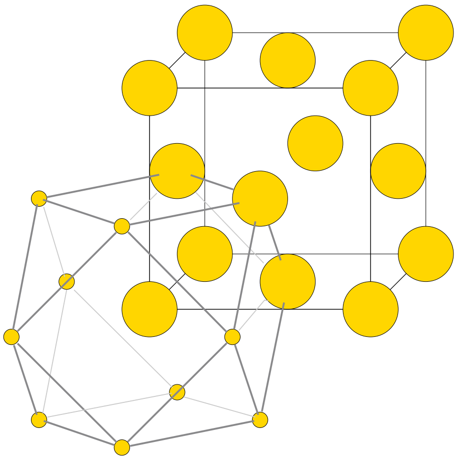


Fig. 12.3: The simple-cubic, the face-centered cubic and the body-centered cubic lattice in the simple cubic unit cell.

12.5 Face centered cubic structure (fcc)

The face-centered cubic structure corresponds to closest sphere packing. There are several structures with closest sphere packing, but the fcc structure is the one with the highest symmetry. The fcc structure is a structural motif, from which we can understand many other structures.

$\vec{T}_1$	0	$\frac{1}{2}a$	$\frac{1}{2}a$
$\vec{T}_2$	$\frac{1}{2}a$	0	$\frac{1}{2}a$
$\vec{T}_3$	$\frac{1}{2}a$	$\frac{1}{2}a$	0
Cu	0	0	0



The fcc-structure is found predominantly for materials without directional bonding. These are, for example, the noble gases in their solid structure. More important are the metals without directional bonding. The red and green elements in Fig. 12.4. It covers most metals except the alkali metals and the transition metals in the middle of the period.

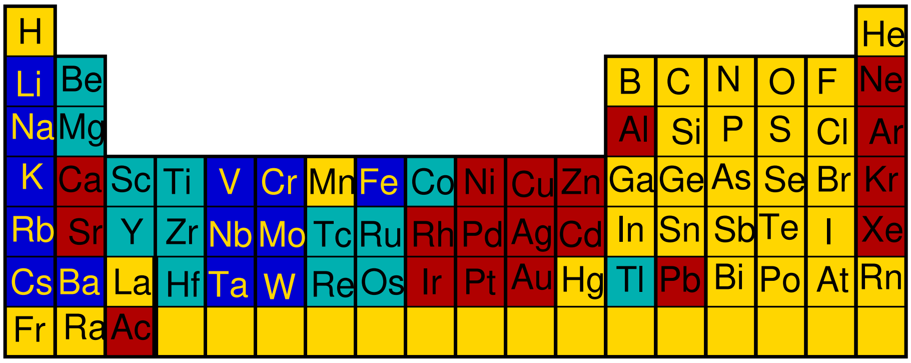

H He
Li Be B C N O F Ne
Na Mg Al Si P S Cl Ar
K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I Xe
Cs Ba La Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn
Fr Ra Ac

Fig. 12.4: Elements with fcc(red), hcp(green) and bcc(blue) structures.

Viewed along the (111) direction we observe in the face-centered cubic lattice three different layers, that can be denoted $\ldots ABCA \ldots$. Each layer consists of equilateral triangles. They differ in their lateral position. The second most important close packing of equally sized spheres is the the hcp structure, where the layering is $\ldots ABAB \ldots$.

Most ionic materials such as oxides exhibit a face-centered cubic arrangement of anions. The anions are much larger than most cations, so that they try to pack as closely as possible. The anions distributed themselves over the interstitial places. The fcc structure contains tetrahedral and octahedral interstitial places. Per fcc site there are two tetrahedral and one octahedral interstitial place. They are located at the positions denoted by O for octahedral and T for tetrahedral in the following table.

$\vec{T}_1$	1	0	0
$\vec{T}_1$	0	1	0
$\vec{T}_1$	0	0	1
atomic site	0	0	0
atomic site	0	$\frac{1}{2}$	$\frac{1}{2}$
atomic site	$\frac{1}{2}$	0	$\frac{1}{2}$
atomic site	$\frac{1}{2}$	$\frac{1}{2}$	0
tetrahedral interstitial	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
tetrahedral interstitial	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
tetrahedral interstitial	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{1}{4}$
tetrahedral interstitial	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{1}{4}$
tetrahedral interstitial	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{3}{4}$
tetrahedral interstitial	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{3}{4}$
tetrahedral interstitial	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{4}$
tetrahedral interstitial	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{3}{4}$
octahedral interstitial	$\frac{1}{2}$	0	0
octahedral interstitial	0	$\frac{1}{2}$	0
octahedral interstitial	0	0	$\frac{1}{2}$
octahedral interstitial	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

The coordination of each atom in the fcc structure is the cubo-octahedron and that of the hcp atom is the twinned cubo-octahedron shown in Fig. 12.2. Each atom has 12 nearest neighbors.

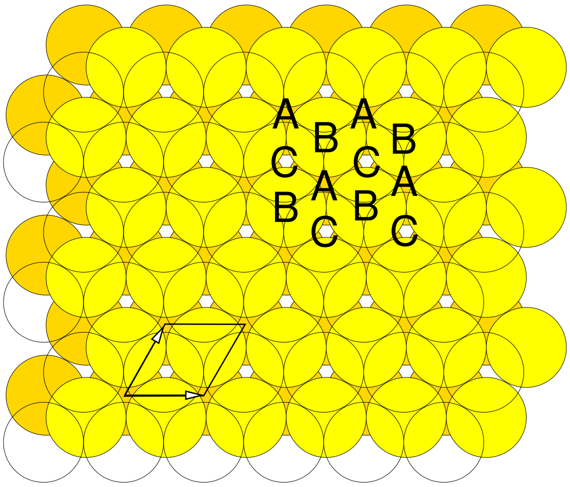


Fig. 12.5: Sphere packing of face-centered cubic (fcc) and hexagonally closed-packed (hcp) structures. The fcc structure, shown, occupies the sites in the sequence ABCABC, while the hexagonally closed packed structure occupies the sites with ABABAB

Elements that crystallize in the fcc structure[33]: Al, Ni, Pd, Ir, Pb,Ca, Cu, Ag, Au, Yb.

12.6 Structure principles derived from closed packing

This is a summary of the structure types and their relation described in more detail below.

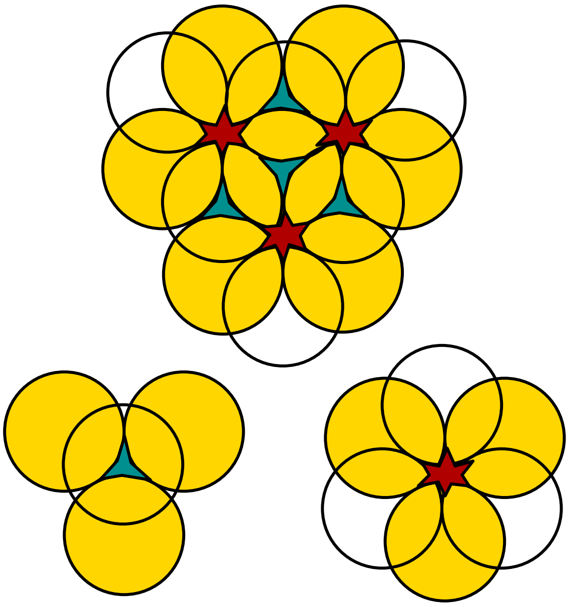


Fig. 12.6: Voids in the fcc structure. The filles circles denote the spheres of the loer layer and the empty circles are the spheres in the upper layer. The red star denotes an octahedral void enclosed by six spheres. The green triangle is a tetrahedral void enclosed by four spheres.

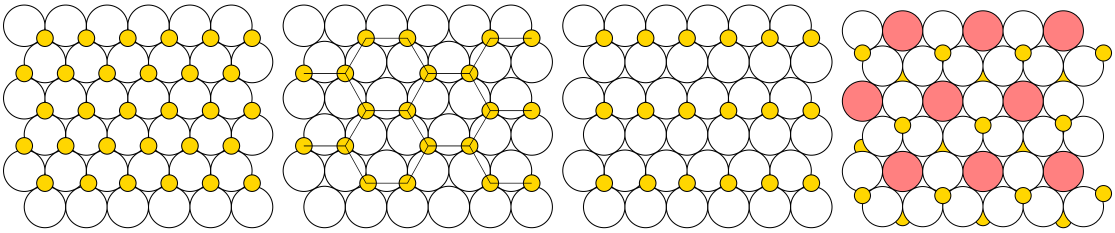


Fig. 12.7: Cation (gold) arrangement of a (111) plane of close-packed anions in octahedral voids. From left to right: Rocksalt structure (NaCl,MgO), corundum structure (Al_2O_3), rutile structure (TiO_2), perovskite ($CaTiO_3$) Calcium ions are drawn in pink. `Editor: Rutile must be checked. Relation to anatase?`

On the first level there are the two closed packed lattices, namely the fcc and the hcp lattice. Each of the lattices has one octahedral void and two tetrahedral voids.

- fcc lattice: Closed sphere packing
  - fill octahedral voids
    - Filling of all octahedral voids of the fcc lattice results in the NaCl structure.
    - Anatase structure (TiO_2) fcc oxygen lattice with Ti in one half of the octahedral holes.
  - fill tetrahedral voids
    - Filling of all tetrahedral voids of the fcc lattice results in the antifluorite (K_2O) structure
    - Filling of one-half of the tetrahedral voids of the fcc lattice results in the zincblende (ZnS) structure.

        - ∗ Bixbyite structure (Mn_2O_3) is obtained by filling 3/4 of the tetrahedral voids. `Editor: Check whether the Bixbyite structure can also be considered as an fcc oxygen lattice with 2/3 of the octahdral voids filled with cations! Reason After a strong relaxation the cations are oictahedrally coordinated.`
    - fill octahedral and tetrahedral voids
        - ∗ Fill all octahedral and all tetrahedral voids of the fcc lattice to obtain the BiF_3 structure.

- hcp lattice: Closed sphere packing. Obtained from fcc by changing the stacking in (111) direction from ABCABC to ABAB.
    - Filling of all octahedral voids of the hcp lattice results in the NiAs structure.
    - Filling of one-half of the tetrahedral of the hcp lattice results in the wurtzite (ZnS) structure
    - Rutile (TiO_2) is a hcp lattice of oxygen ions with one half of the octahedral holes filled by Ti. `http://www.chemtube3d.com/solidstate/_rutile(final).htm`
    - Corundum Al_2O_3 structure is a hcp anion lattice with 2/3 of the octahedral sites occupied by Al. The remaing octahedral voids are arranged to maximize their distance.
- bcc lattice: obtained from fcc by elongation along one of the cubic axes. (Martensitic transformation)

Perovskite (ABO_3)is a defective fcc lattice of oxygen with 1/4 of the oxygen ions replaced by A-type ions. one-quarter of the octahedral voids is occupied by B-type ions. ReO_3 is like a perovskite without A-type ions.

Spinel ($MgAl_2O_4$) face centered oxygen lattice with Mg in the tetrahedral and Al in octahedral voids.

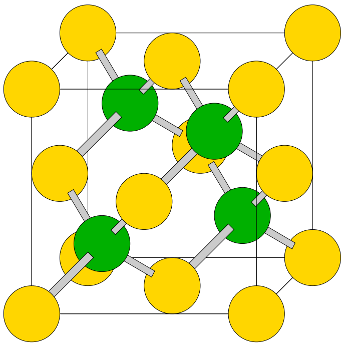
Inverse spinel (Mg_2TiO_4): face centered oxygen lattice with Ti in half of the octahedral voids and Mg in half of the octahedral and all the tetrahedral voids.

12.7 Structures derived from the fcc structure

12.7.1 Zinc-blende structure (ZnS)

If we start from the fcc structure and fill the one-half of the tetrahedral voids with the atoms of a different type we arrive at the zinc-blende structure.

$\vec{T}_1$	0	$\frac{1}{2}a$	$\frac{1}{2}a$
$\vec{T}_2$	$\frac{1}{2}a$	0	$\frac{1}{2}a$
$\vec{T}_3$	$\frac{1}{2}a$	$\frac{1}{2}a$	0
Zn_1	0	0	0
S_2	$\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$



Each atom in the zinc-blende structure is tetrahedrally coordinated by its four nearest neighbors. Thus it is ideal for ions with an average valence of four. Most compound semiconductors relevant for semiconductor industry crystallize in this structure. Compound semiconductors are denoted as IV-IV, III-V, II-VI compounds according to the valency of their constituents.

IV-IV	III-V	II-VI	I-VII
β-SiC	GaSb		CuCl
SiGe	GaAs		
	AlAs		

The zinc-blende structure has large voids, namely in the center of the cube and in the middle of its edges. These are the tetrahedral voids of the fcc structure that have not been filled yet.

The sublattices of both atom types again form two interwoven fcc lattices.

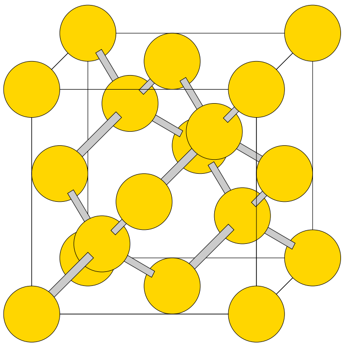
The zincblende structure is identical to the diamond structure, with the only difference that the two atoms are inequivalent. Each atom is tetrahedrally coordinated. Inside the structure large tetrahedral voids are present.

The zinc-blende structure can be seen as an fcc-lattice with one-half of the tetrahedral interstitial sites occupied by cations.

12.7.2 Diamond structure

A simpler variant of the zinc-blende structure is the diamond structure. The only difference to the zinc-blende structure is that all atoms are identical in the diamond structure.

$\vec{T}_1$	0	$\frac{1}{2}a$	$\frac{1}{2}a$
$\vec{T}_2$	$\frac{1}{2}a$	0	$\frac{1}{2}a$
$\vec{T}_3$	$\frac{1}{2}a$	$\frac{1}{2}a$	0
C_1	0	0	0
C_2	$\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$



Each atom is tetrahedrally coordinated. Inside the structure large tetrahedral voids are present, namely just the remaining tetrahedral voids of the fcc structure that has not been filled. tetrahedral voids occupied.

Note the zig-zag chains running along the (110) directions.

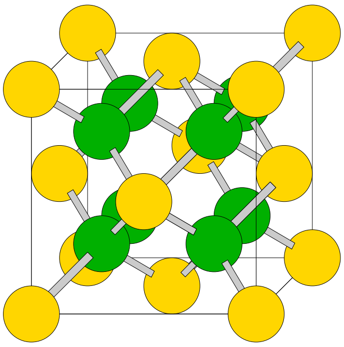
Both carbon and silicon crystallize in the diamond structure.

12.7.3 Fluorite structure (CaF_2)

If we fill all the tetrahedral voids of the fcc structure with ions of a different type, we arrive at the fluorite structure. The formula unit for the fluorite structure is AB_2.

Each fluorine is tetrahedrally coordinated with Ca atoms. Each calcium atom is centered in a cube of eight fluorine atoms.Thus the calcium is has the same environment as an atom in the bcc structure. The fluorine ions form a simple-cubic structure among each other, with every second center occupied by a calcium atom. Thus the fluorite structure is related to the CsCl structure such that only half of the cation positions of the CsCl structure are occupied.

$\vec{T}_1$	0	$\frac{1}{2}a$	$\frac{1}{2}a$
$\vec{T}_2$	$\frac{1}{2}a$	0	$\frac{1}{2}a$
$\vec{T}_3$	$\frac{1}{2}a$	$\frac{1}{2}a$	0
Ca_1	0	0	0
F_1	$\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$
F_2	$\frac{1}{4}a$	$\frac{1}{4}a$	$-\frac{1}{4}a$



Cubic unit cell (4 formula units)

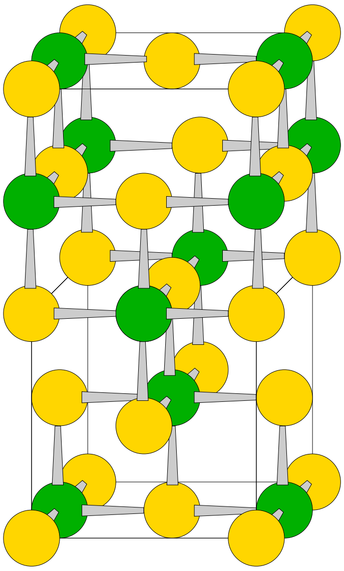
$\vec{T}_1$	a	0	0
$\vec{T}_2$	0	a	0
$\vec{T}_3$	0	0	a
Ca_1	0	0	0
Ca_2	$\frac{1}{2}a$	$\frac{1}{2}a$	0
Ca_3	0	$\frac{1}{2}a$	$\frac{1}{2}a$
Ca_4	$\frac{1}{2}a$	0	$\frac{1}{2}a$
F_1	$\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$
F_2	$\frac{1}{4}a$	$\frac{3}{4}a$	$\frac{1}{4}a$
F_3	$\frac{3}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$
F_4	$\frac{3}{4}a$	$\frac{3}{4}a$	$\frac{1}{4}a$
F_5	$\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{3}{4}a$
F_6	$\frac{1}{4}a$	$\frac{3}{4}a$	$\frac{3}{4}a$
F_7	$\frac{3}{4}a$	$\frac{1}{4}a$	$\frac{3}{4}a$
F_8	$\frac{3}{4}a$	$\frac{3}{4}a$	$\frac{3}{4}a$

12.7.4 Anti-fluorite structure

The antifluorite structure is identical to the fluorite structure with the main difference that the anions and cations are exchanged. While the fcc positions are occupied by cations in the fluorite structure by cations, they are occupied by anions in the anti-fluorite structure.

12.7.5 Anatase structure (TiO_2)

Anatase is a metastable form of TiO_2. This structure is derived from an fcc-type anion lattice, with one-half of the octahedral voids occupied by cations. For the sake of clarity we only show the idealized positions of anatase



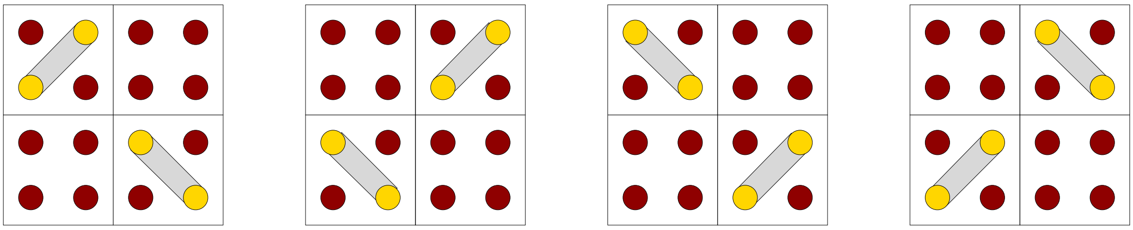
Each oxygen ion is three-fold coordinated by cation. This rule can be extracted from the composition and the octahedral environment of the cations. Because there are six bonds from each cation to the anion neighbors, this is also the number of bonds from each anion to the cation per cation. Per anion there are half as many bonds, namely three.

The ideal structure shown above undergoes relaxations, which distorts the octahedra and makes the angles on the three-fold coordinated oxygen ions more equal.

Anatase is the fcc-type variant of the rutile structure, which is based on an hcp-type anion lattice.

12.7.6 Bixbyite structure (Mn_2O_3)

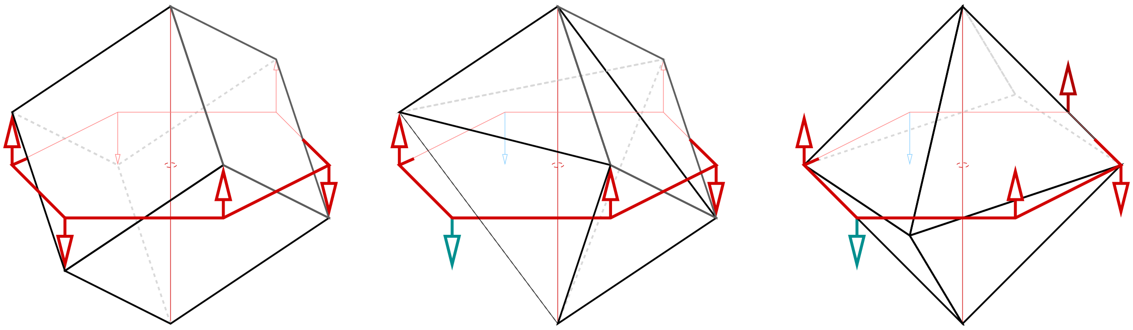
The bixbyite structure is a variant of the fluorite structure with a quarter of oxygen atoms removed. The distribution of the oxygen vacancies is similar to sudoku game. In each of the three cartesian directions there shall be three oxygen atoms in a row, separated by a vacancy. The pattern of oxygen vacancies is shown in the following figure.



Only the oxygen sublattice is shown.

The Mn-ions among themselfes form an fcc-lattice.

Each Mn atom is in the center of a cube of oxygen atoms of which two are removed. Thus there are 6 neighbors. This environment PROBABLY arranges itself into the shape of an octahedron. The process of this rearrangement is shown here.



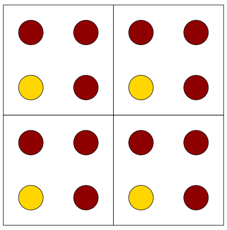
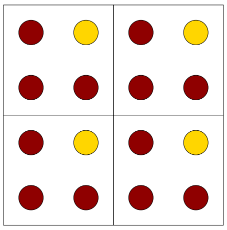
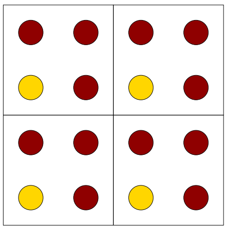
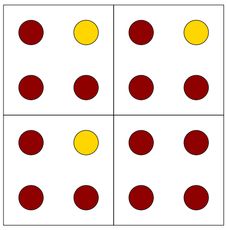
First two of the corners are removed. Then the four remaining equatorial atoms are shifted into their new positions.

The following table provides the idealized bixbyite structure expressed by the ideal lattice points of the fluorite lattice. This lattice undergoes severe rearrangements until the true bixbyite structure is reached.

$\vec{T}_1$	$-a$	a	a
$\vec{T}_2$	a	$-a$	a
$\vec{T}_3$	a	a	$-a$
Mn_1	0	0	0
Mn_2	$\frac{1}{2}a$	$\frac{1}{2}a$	0
Mn_3	a	0	0
Mn_4	$\frac{3}{2}a$	$\frac{1}{2}a$	0
Mn_5	0	a	0
Mn_6	$\frac{1}{2}a$	$\frac{3}{2}a$	0
Mn_7	a	a	0
Mn_8	$\frac{3}{2}a$	$\frac{3}{2}a$	0
Mn_9	$\frac{1}{2}a$	0	$\frac{1}{2}a$
Mn_{10}	$\frac{3}{2}a$	0	$\frac{1}{2}a$
Mn_{11}	0	$\frac{1}{2}a$	$\frac{1}{2}a$
Mn_{12}	a	$\frac{1}{2}a$	$\frac{1}{2}a$
Mn_{13}	$\frac{1}{2}a$	a	$\frac{1}{2}a$
Mn_{14}	$\frac{3}{2}a$	a	$\frac{1}{2}a$
Mn_{15}	0	$\frac{3}{2}a$	$\frac{1}{2}a$
Mn_{16}	a	$\frac{3}{2}a$	$\frac{1}{2}a$
V_1	$\frac{1}{4}a$	$\frac{3}{4}a$	$\frac{1}{4}a$
V_2	$\frac{3}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$
V_3	$1\frac{1}{4}a$	$1\frac{1}{4}a$	$\frac{1}{4}a$
V_4	$1\frac{3}{4}a$	$1\frac{3}{4}a$	$\frac{1}{4}a$
V_5	$1\frac{1}{4}a$	$\frac{3}{4}a$	$\frac{3}{4}a$
V_6	$1\frac{3}{4}a$	$\frac{1}{4}a$	$\frac{3}{4}a$
V_7	$\frac{1}{4}a$	$1\frac{1}{4}a$	$\frac{3}{4}a$
V_8	$\frac{3}{4}a$	$1\frac{3}{4}a$	$\frac{3}{4}a$

O_1	$\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$
O_2	$\frac{3}{4}a$	$\frac{3}{4}a$	$\frac{1}{4}a$
O_3	$1\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$
O_4	$1\frac{1}{4}a$	$\frac{3}{4}a$	$\frac{1}{4}a$
O_5	$1\frac{3}{4}a$	$\frac{1}{4}a$	$\frac{1}{4}a$
O_6	$1\frac{3}{4}a$	$\frac{3}{4}a$	$\frac{1}{4}a$
O_7	$\frac{1}{4}a$	$1\frac{1}{4}a$	$\frac{1}{4}a$
O_8	$\frac{1}{4}a$	$1\frac{3}{4}a$	$\frac{1}{4}a$
O_9	$\frac{3}{4}a$	$1\frac{1}{4}a$	$\frac{1}{4}a$
O_{10}	$\frac{3}{4}a$	$1\frac{3}{4}a$	$\frac{1}{4}a$
O_{11}	$1\frac{1}{4}a$	$1\frac{3}{4}a$	$\frac{1}{4}a$
O_{12}	$1\frac{3}{4}a$	$1\frac{1}{4}a$	$\frac{1}{4}a$
O_{13}	$\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{3}{4}a$
O_{14}	$\frac{1}{4}a$	$\frac{3}{4}a$	$\frac{3}{4}a$
O_{15}	$\frac{3}{4}a$	$\frac{1}{4}a$	$\frac{3}{4}a$
O_{16}	$\frac{3}{4}a$	$\frac{3}{4}a$	$\frac{3}{4}a$
O_{17}	$1\frac{1}{4}a$	$\frac{1}{4}a$	$\frac{3}{4}a$
O_{18}	$1\frac{3}{4}a$	$\frac{3}{4}a$	$\frac{3}{4}a$
O_{19}	$\frac{1}{4}a$	$1\frac{3}{4}a$	$\frac{3}{4}a$
O_{20}	$\frac{3}{4}a$	$1\frac{1}{4}a$	$\frac{3}{4}a$
O_{21}	$1\frac{1}{4}a$	$1\frac{1}{4}a$	$\frac{3}{4}a$
O_{22}	$1\frac{1}{4}a$	$1\frac{3}{4}a$	$\frac{3}{4}a$
O_{23}	$1\frac{3}{4}a$	$1\frac{1}{4}a$	$\frac{3}{4}a$
O_{24}	$1\frac{3}{4}a$	$1\frac{3}{4}a$	$\frac{3}{4}a$

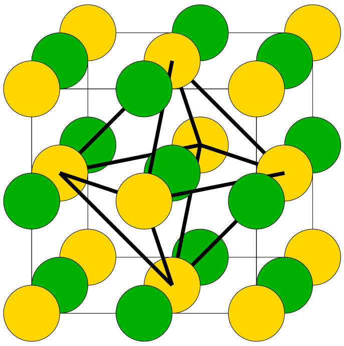
Another ordering of the oxygen vacancies is the following, which however does not lead to the bixbyite structure.

12.7.7 Rock salt (NaCl)

The rock salt structure is obtained from the fcc structure by occupying all octahedral voids with cations. Each ion is octahedrally coordinated by the opposite species. Thus the lattice is composed of face-sharing octahedra with anions at the corners and cations in their center, or vice versa.

$\vec{T}_1$	0	$\frac{1}{2}a$	$\frac{1}{2}a$
$\vec{T}_2$	$\frac{1}{2}a$	0	$\frac{1}{2}a$
$\vec{T}_3$	$\frac{1}{2}a$	$\frac{1}{2}a$	0
Na_1	0	0	0
Cl_1	$\frac{1}{2}a$	0	0



The rock salt structure is a simple cubic structure with the positions alternatively occupied by cations and anions.

Substances

	a	ref
NaCl	5.63Å	Kittel
KCl	6.29Å	Kittel
KBr	6.59Å	Kittel
MgO	4.20Å	Kittel
MnO	4.43Å	Kittel
PbS	5.92Å	Kittel
LiH	4.08Å	Kittel

12.7.8 Spinel structure ($MgAl_2O_4$)

Spinel is not finished

The cubic unit cell contains eight formula units.

Voids of the O-lattice The spinel structure is based on an fcc lattice of oxygen ions. One-half of the octahedral voids are occupied by Al^{3+} ions and 1/8 of the tetrahedral voids are filled with Mg^{2+} ions.

Network of octahedra

1. Let me start from the fcc anion lattice in the cubic unit cell with eight formula units.

2. In the first (001) plane with $z = 0$, fill every second row of octahedral voids by Al. Together with the apical oxygen ions of the neighboring plane, one-dimensional ribbons of edge-connected octahedra are formed that run along the (110) direction. They are separated by a chain of vacant octahedral voids.
3. Similarly, in third (001) plane ($z = 1$), fill every second row of octahedral voids by Al. This forms again ribbons of octahedra running along (110). They are horizontally displaced relative to the first layer, so that their distance is maximized
4. In the odd-numbered planes $z = 0.5$ and $z = 1.5$, fill one-half of the octahedral sites, but now orient the ribbons orthogonal to the ones in the even-numbered planes, i.e. along $(1\bar{1}0)$. As for the even-nombered planes the ribbons are stacked to maximize their distance.
5. Fill tetrahedral voids in between the oxygen planes with Mg ions. Per layer in the unit cell two Mg sites are introduced. The orthogonal ribbons of the octaheder chains in the layer above and below form the boundary of squares. The tetralhedral void in the center of the square is filled with the Mg ion.

Each octahedron is edge-connected with the neighboring octahedra in the ribbons. Furthermore it is edge connected with with two octahedra each from the neighboring oxygen planes. As a result, each octahedron is edge connected to six further octahdera. For each octahedron, two opposite faces of the octahedron are edge connected to three other octahedra each.

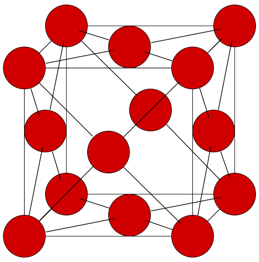
The following page has useful images of the atomic positions in the spinel structure: `https://www.tau.ac.il/~morr`

Table 12.1: Idealized spinel structure Al_2MgO_4 in the non-primitive cubic unit cell.

$\vec{T}_1$	1	1	0	spinel $(Al_2MgO_4)_4$ 28 atoms											
$\vec{T}_2$	1	−1	0												
$\vec{T}_3$	0	0	2												
O	0	0	0	O	$\frac{1}{2}$	0	$\frac{1}{2}$	O	0	0	0	O	$\frac{1}{2}$	0	$\frac{3}{2}$
O	$\frac{1}{2}$	$\frac{1}{2}$	0	O	1	$\frac{1}{2}$	$\frac{1}{2}$	O	$\frac{1}{2}$	$\frac{1}{2}$	1	O	1	$\frac{1}{2}$	$\frac{3}{2}$
O	$\frac{1}{2}$	$-\frac{1}{2}$	0	O	1	$-\frac{1}{2}$	$\frac{1}{2}$	O	$\frac{1}{2}$	$-\frac{1}{2}$	1	O	1	$-\frac{1}{2}$	$\frac{3}{2}$
O	1	1	0	O	$\frac{3}{2}$	0	$\frac{1}{2}$	O	1	1	0	O	$\frac{3}{2}$	0	$\frac{3}{2}$
Al	$\frac{1}{2}$	0	0	Al	0	0	$\frac{1}{2}$	Al	1	$-\frac{1}{2}$	1	Al	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$
Al	1	$\frac{1}{2}$	0	Al	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	Al	$\frac{3}{2}$	0	1	Al	1	0	$\frac{3}{2}$
Mg	$\frac{5}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	Mg	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{3}{4}$	Mg	$\frac{5}{4}$	$\frac{3}{4}$	$\frac{5}{4}$	Mg	$\frac{3}{4}$	$-\frac{3}{4}$	$\frac{7}{4}$

The idealized atomic positions of the spinel structure in the 56-atom cubic $(Al_2MgO_4)_8$ unit cell are provided in table 12.2. The atomic structure in a format suitable for as input for the CP-PAW code is given in appendix E.7 on p. 194. The atomic positions are also given in a smaller 28-atom unit cell in table 12.1.

`Editor: The following is not completed. Complete or remove!`

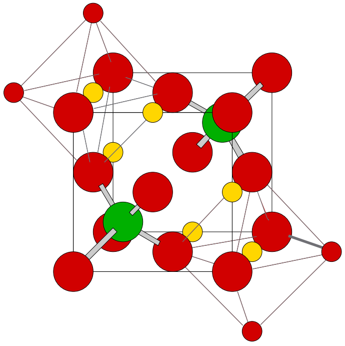


In the first cubic cell two of the tetrahedral voids are occupied by

Table 12.2: Idealized spinel structure $(Al_2MgO_4)_8$ in the non-primitive cubic unit cell. The positions for each element are arranged line-by-line and plane-by-plane. The atomic structure in a format suitable for as input for the CP-PAW code is given in appendix E.7 on p. 194.

$\vec{T}_1$	2	0	0	spinel cubic 56 atoms											
$\vec{T}_2$	0	2	0												
$\vec{T}_3$	0	0	2												
O	0	0	0	O	$\frac{1}{2}$	0	$\frac{1}{2}$	O	0	0	1	O	$\frac{1}{2}$	0	$\frac{3}{2}$
O	1	0	0	O	$\frac{3}{2}$	0	$\frac{1}{2}$	O	1	0	1	O	$\frac{3}{2}$	0	$\frac{3}{2}$
O	$\frac{1}{2}$	$\frac{1}{2}$	0	O	0	$\frac{1}{2}$	$\frac{1}{2}$	O	$\frac{1}{2}$	$\frac{1}{2}$	1	O	0	$\frac{1}{2}$	$\frac{3}{2}$
O	$\frac{3}{2}$	$\frac{1}{2}$	0	O	1	$\frac{1}{2}$	$\frac{1}{2}$	O	$\frac{3}{2}$	$\frac{1}{2}$	1	O	1	$\frac{1}{2}$	$\frac{3}{2}$
O	0	1	0	O	$\frac{1}{2}$	1	$\frac{1}{2}$	O	0	1	1	O	$\frac{1}{2}$	1	$\frac{3}{2}$
O	1	1	0	O	$\frac{3}{2}$	1	$\frac{1}{2}$	O	1	1	1	O	$\frac{3}{2}$	1	$\frac{3}{2}$
O	$\frac{1}{2}$	$\frac{3}{2}$	0	O	0	$\frac{3}{2}$	$\frac{1}{2}$	O	$\frac{1}{2}$	$\frac{3}{2}$	1	O	0	$\frac{3}{2}$	$\frac{3}{2}$
O	$\frac{3}{2}$	$\frac{3}{2}$	0	O	1	$\frac{3}{2}$	$\frac{1}{2}$	O	$\frac{3}{2}$	$\frac{3}{2}$	1	O	1	$\frac{3}{2}$	$\frac{3}{2}$
Al	$\frac{1}{2}$	0	0	Al	1	0	$\frac{1}{2}$	Al	$\frac{3}{2}$	0	1	Al	0	0	$\frac{3}{2}$
Al	1	$\frac{1}{2}$	0	Al	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	Al	0	$\frac{1}{2}$	1	Al	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{3}{2}$
Al	$\frac{3}{2}$	1	0	Al	0	1	$\frac{1}{2}$	Al	$\frac{1}{2}$	1	1	Al	1	1	$\frac{3}{2}$
Al	0	$\frac{3}{2}$	0	Al	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	Al	1	$\frac{3}{2}$	1	Al	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$
Mg	$\frac{7}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	Mg	$\frac{5}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	Mg	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{5}{4}$	Mg	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{7}{4}$
Mg	$\frac{3}{4}$	$\frac{5}{4}$	$\frac{1}{4}$	Mg	$\frac{1}{4}$	$\frac{7}{4}$	$\frac{3}{4}$	Mg	$\frac{7}{4}$	$\frac{5}{4}$	$\frac{5}{4}$	Mg	$\frac{5}{4}$	$\frac{7}{4}$	$\frac{7}{4}$

Mg atoms. They are located as far from each other on a main diagonal of the cube. Now we occupy the octahedral voids of the fcc structure, which are farthest from the tetrahedra, with Al ions. This means that octahedra that share faces with the tetrahedra are avoided.



12.7.9 ABC-Birnessite

Birnessite is a layered calcium-manganite. The structure here is not confirmed. It is an idealized model, which probably reflects the main characteristics of Birnessite.

The basic building block of birnessite is a layer of edge-sharing MnO_2 octahedra. Each such layer can be considered as two close packed oxygen planes with Mn on every other octahedral interstitial site, in between the oxygen layers. The MnO_2 layers are saturated on the top and the bottom by Ca ions. These Ca-covered layers are sparated by one monolayer of water molecules.

This material can also be regarded as a close-packed lattice of oxygen ions with Ca-ions and Mn-ions on octahedral positions. One of oxygen layers is made of water. Thus it has the stacking

$$\mathrm{Ca}^{2+}\text{-}\Big[\mathrm{H_2O}\text{-}\mathrm{Ca}^{2+}\text{-}\mathrm{O}^{2-}\text{-}\mathrm{Mn}^{4+}\text{-}\mathrm{O}^{2-}\text{-}\mathrm{Ca}^{2+}\Big]\text{-}\mathrm{H_2O}\text{-}$$

$\vec{T}_1$	1	0	0
$\vec{T}_2$	$\frac{1}{2}$	$\frac{1}{2}\sqrt{3}$	0
$\vec{T}_3$	0	0	$3\sqrt{\frac{2}{3}}$
O_H	0	0	0
H_1	x	0	0
H_2	0	x	0
Ca_l	1	$\frac{1}{3}\sqrt{3}$	$\frac{1}{2}\sqrt{\frac{2}{3}}$
O_l	$\frac{1}{2}$	$\frac{1}{6}\sqrt{3}$	$\sqrt{\frac{2}{3}}$
Mn	0	0	$\frac{3}{2}\sqrt{\frac{2}{3}}$
O_u	1	$\frac{1}{3}\sqrt{3}$	$2\sqrt{\frac{2}{3}}$
Ca_u	$\frac{1}{2}$	$\frac{1}{6}\sqrt{3}$	$\frac{5}{2}\sqrt{\frac{2}{3}}$

12.8 Hexagonally closed packed structure (hcp)

$\vec{T}_1$	a	$0a$	0
$\vec{T}_2$	$\frac{1}{2}a$	$\frac{1}{2}\sqrt{3}a$	0
$\vec{T}_3$	0	0	$2\sqrt{\frac{2}{3}}a$
A	0	0	0
B	$\frac{1}{2}a$	$\frac{1}{6}\sqrt{3}$	$\sqrt{\frac{2}{3}}a$

The coordination polyhedron of the hexagonally closed packed structure is the twinned cubo-octahedron. The face-centered cubic and the hexagonally closed packed structures correspond to a closed packed sphere packing.

Also the hexagonally closed packed structure contains two tetrahedral interstitial sites and one octahedral site per hcp site. The are located at

tetrahedral interstitial	$\frac{1}{2}a$	$\frac{1}{6}\sqrt{3}$	$\frac{1}{4}\sqrt{\frac{2}{3}}a$
tetrahedral interstitial	0	0	$\frac{3}{4}\sqrt{\frac{2}{3}}a$
tetrahedral interstitial	0	0	$\frac{5}{4}\sqrt{\frac{2}{3}}a$
tetrahedral interstitial	$\frac{1}{2}a$	$\frac{1}{6}\sqrt{3}$	$\frac{7}{4}\sqrt{\frac{2}{3}}a$
octahedral interstitial	a	$\frac{2}{6}\sqrt{3}$	$\frac{1}{2}\sqrt{\frac{2}{3}}a$
octahedral interstitial	a	$\frac{2}{6}\sqrt{3}$	$\frac{3}{2}\sqrt{\frac{2}{3}}a$

Elements that crystallize in the hexagonally closed packed structure[33]: Be, Ti, Zn, Tl, Ho, Mg, Co, Zr, Gd.

12.9 Structures derived from the hcp structure

12.9.1 Graphite

$\vec{T}_1$	a	0	0
$\vec{T}_2$	$\frac{1}{2}a$	$\frac{1}{2}\sqrt{3}$	0
$\vec{T}_3$	0	0	c
C_1	0	0	0
C_2	$\frac{1}{2}a$	$\frac{1}{2\sqrt{3}}a$	0
C_3	0	0	$\frac{1}{2}c$
C_4	a	$\frac{1}{\sqrt{3}}a$	$\frac{1}{2}c$

Each atom is coordinated to three neighbors in the plane. The planes is a honeycomb lattice of six-fold rings.

Substances

	a	ref
C	??Å	??
Si	??Å	??
Ge	??Å	??

12.9.2 NiAs structure (B8-type)

The NiAs structure can be considered the hexagonal variant of the rock-salt structure. Here the anions (As) form a hexagonally closed packed lattice, and all octahedral voids are filled by nickel ions.[2]

The nickel sublattice consists of a closed-packed layer that is stacked ontop of each other in an AAA-stacking.

$\vec{T}_1$	a	0	0
$\vec{T}_2$	$\frac{1}{2}a$	$\frac{1}{2}\sqrt{3}$	0
$\vec{T}_3$	0	0	c
Ni_1	0	0	0
Ni_2	0	0	$\frac{1}{2}c$
As_1	$\frac{1}{2}a$	$\frac{1}{2\sqrt{3}}a$	$\frac{1}{4}c$
As_2	a	$\frac{1}{\sqrt{3}}a$	$\frac{3}{4}c$

	a	c	ref
NiAs	3.618Å	5.034 Å	??

12.9.3 α-Corundum

The corundum structure can be seen as a distorted hcp lattice of oxygen atoms with cations filling two-thirds of the octahedral interstitial sites. The cations fill in each row within the plane two positions and then leave one vacant. The arrangement of Al cations and vacant octahedral sites is such that the Al atoms form condensed six-fold rings similar to the graphite structure, where each ring encloses a vacant octahedral site.

---

[2] In contrast, the anions (Cl) in the rocksalt (NaCl) structure form an cubic closed packed lattice (fcc) and while the cations (Na) fill all octahedral voids.

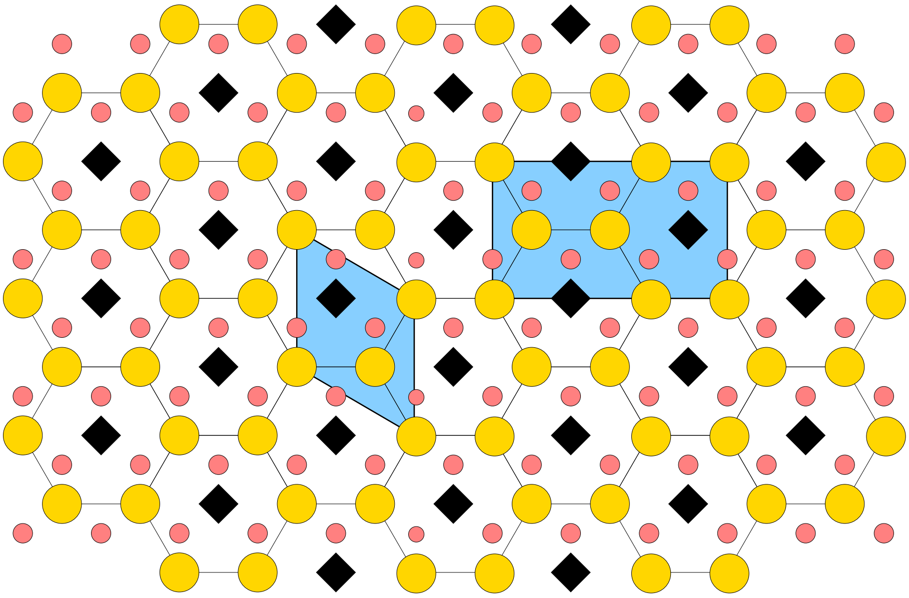


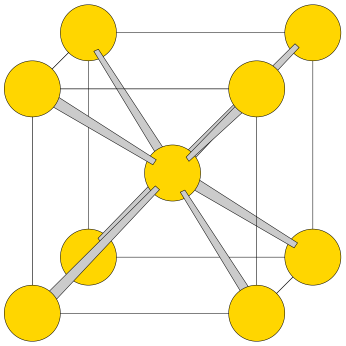
Fig. 12.8: Arrangement of Al ions in the α-corundum structure. Aluminium atoms are shown in gold, and the octahedral interstices as black squares. The oxygen atoms in the layer below in red.

12.9.4 Cadmium iodide

`http://en.wikipedia.org/wiki/Cadmium_iodide`

12.10 Body-centered cubic structure

$\vec{T}_1$	$-\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$
$\vec{T}_2$	$\frac{1}{2}a$	$-\frac{1}{2}a$	$\frac{1}{2}a$
$\vec{T}_3$	$\frac{1}{2}a$	$\frac{1}{2}a$	$-\frac{1}{2}a$
Fe	0	0	0



The bcc structure can be obtained from the fcc structure via the so-called **martensitic transition**, which plays an important role in steel. Iron exists in two stable forms, as the more stable, ferromagnetic ferrite or as the non-magnetic austenite form. Martensite is a very brittle and hard phase of iron containing carbon, which is the main component of steel. It has a body-centered tetragonal structure. It is obtained by cooling the high-temperature form of carbon-containing iron, austenite. In order to obtain the bcc structure we compress the fcc in z-direction and expand it simultaneously in x and y directions.

Substances

Body centered cubic crystals are usually found in the middle of the transition metal series. This is because the structure exhibits stronger bonding in the d-shell than the closed packed structures.

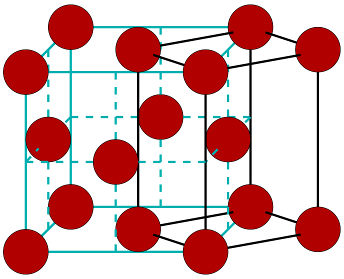


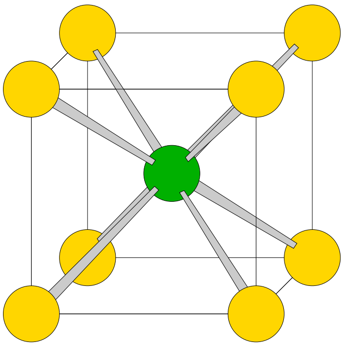
Fig. 12.9: Sketch of the martensitic transformation. Shown is the fcc lattice with the distorted fcc cell inscribed in black. By compressing the vertical direction and expanding the horizontal directions, the fcc lattice can be transformed continuously into the bcc lattice.

	a	ref
Li	3.491Å	Kittel
Na	4.225Å	Kittel
K	5.225Å	Kittel
Rb	5.585Å	Kittel
Cs	6.045Å	Kittel
Fe	2.87Å	Kittel
Cr	2.88Å	Kittel
V	3.03Å	Kittel
Nb	3.30Å	Kittel
Mo	3.15Å	Kittel
Ta	3.30Å	Kittel
W	3.16Å	Kittel

12.11 Structures derived from the bcc structure

12.11.1 Caesium chloride structure (CsCl)

$\vec{T}_1$	a	0	0
$\vec{T}_2$	0	a	0
$\vec{T}_3$	0	0	a
Cs_1	0	0	0
Cl_1	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$



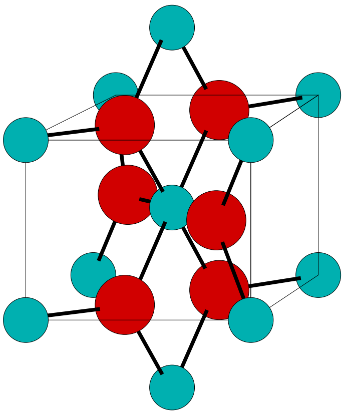
The structure is a body-centered cubic structure with the positions alternatively occupied by cations and anions.

Substances

	a	ref
CsCl	??Å	??

12.11.2 Rutile (C4-type)

$\vec{T}_1$	a	0	0
$\vec{T}_2$	0	a	0
$\vec{T}_3$	0	0	c
Ti_1	0	0	0
Ti_2	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}c$
O_1	$\frac{1}{3}a$	$\frac{1}{3}a$	0
O_2	$\frac{2}{3}a$	$\frac{2}{3}a$	0
O_3	$\frac{1}{3}a$	$\frac{2}{3}a$	$\frac{1}{2}c$
O_4	$\frac{2}{3}a$	$\frac{1}{3}a$	$\frac{1}{2}c$



The oxygen atoms are three-fold coordinated. The Ti atoms are six-fold coordinated to oxygen atoms. The Ti atoms occupy the body centered lattice positions in the tetragonal unit cell.

The ideal c/a ratio is $\sqrt{\frac{2}{3}} = 0.816497$. It is obtained from the condition that each oxygen atom is exactly trigonally coordinated.

The rutile structure is formed from chains of edge-connected TiO_6 octhahedra, which are connected such that the unbridged corners bind to a bridging corner of a neighboring chain. This explains the three-fold coordination of the oxygen atoms.

Questionable! Check the following: The rutile structure is a hcp lattice of oxygen atoms with anions filling the octahedral voids.

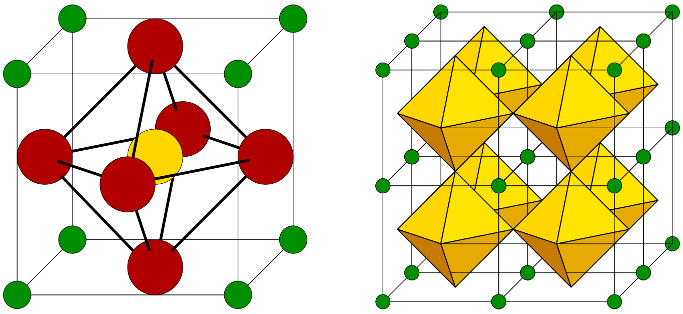
Substances

	a	c	ref
FeF_2	4.6966Å	3.3091 Å	??
TiO_2	??Å	?? Å	??

FeF_2 is anti-ferromagnetic with spins aligned along the c-axis.

12.12 Perovskite

$\vec{T}_1$	a	0	0
$\vec{T}_2$	0	a	0
$\vec{T}_3$	0	0	a
Sr_1	0	0	0
Ti_1	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$
O_1	0	$\frac{1}{2}a$	$\frac{1}{2}a$
O_2	$\frac{1}{2}a$	0	$\frac{1}{2}a$
O_3	$\frac{1}{2}a$	$\frac{1}{2}a$	0



The ideal perovskite lattice can be derived from an fcc-lattice of oxygen ions: In the most simple case, one introduces positive charges on a bcc-type lattice, half of which replace an oxygen ion while the other half is inserted into an octahedral void of the oxygen lattice. The ions replacing the oxygen ion form a simple cubic lattice and are shown as green balls in the figure above. They are called A-type ions, referring to the general formula unit ABO_3 of a perovskite. The other ion, called the B-type ion, occupy the center of the cube formed by the A-type ions. Thus, the AO_3 sublattice forms a fcc lattice, while the Ti-atoms are located in its octahedral voids.

The lattice can also be seen as a corner-sharing array of BO_6 (TiO_6) octahedra, with the A-type ions (Sr) atoms located in the voids between eight such octahedra.

The Sr sees the environment of the fcc lattice with twelve neighbors, that is a cubo-octahedron. The oxygen atoms are two-fold coordinated. The Ti atoms are octahedrally coordinated to oxygen atoms.

The members of the perovskite family rarely have the ideal cubic structure. In fact there is a large zoo of structural distortions and electronic instabilities. This in turn causes a tremendous richness in physical properties such as insulators, metals, ferroelectrics, magnetism, giant magnetoresistance, superconductivity, and many more. We will discuss them in little more detail in section 12.13

The most prominent distortions are

- tiltings of the BO_6 octahedra
- Jahn-Teller distortions of the BO_6 octahedra
- ferroelectric displacements of the BO_6 octahedra

An excellent description of the perovskite family is given in the Thesis of F. Lindberg[38].

The magnetic arrangements are classified according to Wollan[39] as shown in Fig. 12.10

Mor information can be found in appendix **??**.

Substances

	a	ref
$SrTiO_3$	??Å	??
$LaAlO_3$	??Å	??

12.12.1 YBCO

The high-T_c Superconductors are derived from the perovskite structure. The prototypical structure is yttrium Barium Copper oxide $Y_1Ba_2Cu_3O_7$. commonly termed “123”compound. The YBCO has a transition temperature of 77 K.

Let us start[**?**] the mental construction from a fictitious perovskite $BaCuO_3$, where the copper lies in the center of the octahedron. The structure of YBCO is obtained by tripling the perovskite structure in c direction and obtain a composition $Ba_3Cu_3O_9$. One of the Ba-atoms is replaced by Yttrium, which yields $Y_1Ba_2Cu_3O_9$. Finally we introduce two oxygen vacancies, namely one in the Y-layer and one in the Cu layer farthest from the Y layer. Thus the layer composition is

$$\ldots\text{-BaO-CuOV}_O\text{-BaO-CuO}_2\text{-YV}_O\text{-CuO}_2\text{-BaO-CuOV}_O\text{-BaO-}\ldots$$

With V_O we denote an oxygen vacancy.

h!

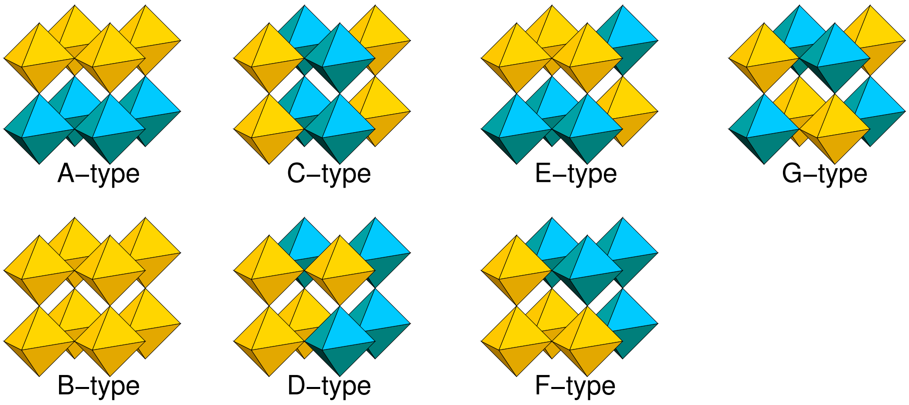
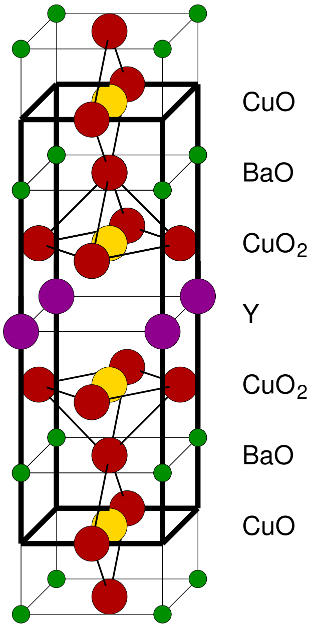


Fig. 12.10: Classification of magnetic orderings of perovskites according to Wollan[39]. One octant of the magnetic unit cell, that is all the inequivalent octahedra, is shown. The gold and blue octahedra represent central atoms with up and down spin arrangement.

$\vec{T}_1$	a	0	0
$\vec{T}_2$	0	a	0
$\vec{T}_3$	0	0	3a
Cu_1	$\frac{1}{2}a$	$\frac{1}{2}a$	0
O_1	$\frac{1}{2}a$	0	0
Ba_1	0	0	$\frac{1}{2}a$
O_2	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$
Cu_2	$\frac{1}{2}a$	$\frac{1}{2}a$	a
O_3	$\frac{1}{2}a$	0	a
O_4	0	$\frac{1}{2}a$	a
Y_1	0	0	$\frac{3}{2}a$
Cu_3	$\frac{1}{2}a$	$\frac{1}{2}a$	2a
O_5	$\frac{1}{2}a$	0	2a
O_6	0	$\frac{1}{2}a$	2a
Ba_2	0	0	$\frac{5}{2}a$
O_7	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{5}{2}a$



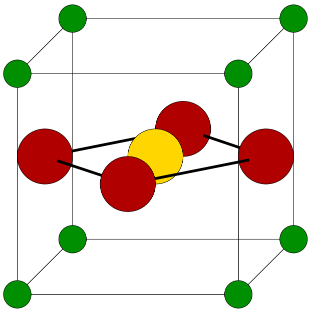
The important structural element of the YBCO structure are the Cu atoms in the square planar environment of oxygen ligands. The CuO_2 layers form square networks of Cu atoms with O-atoms on the connecting lines. The CuO planes with the oxygen on-top and below are corner sharing CuO_4-squares.

The charge distribution of the layers is such that we can assume that the Cu atoms are formally in the +2 charge state, but there is one hole h^+ per unit cell, because the trivalent Y^{3+} cannot compensate the charge of 4− from the neighboring CuO_2 layers.

12.12.2 $CaCuO_2$

The $CaCuO_2$ structure is the prototype structure containing the square CuO_2 lattice present in the high-T_c superconductors.

$\vec{T}_1$	a	0	0
$\vec{T}_2$	0	a	0
$\vec{T}_3$	0	0	a
Ca_1	0	0	0
Cu_1	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$
O_1	$\frac{1}{2}a$	0	$\frac{1}{2}a$
O_2	0	$\frac{1}{2}a$	$\frac{1}{2}a$



12.13 Mixture of structural principles

The perovskites family

Most perovskites do not have the ideal cubic structure described in section 12.12, but they undergo a range of distortions. These distortions can be classified as Jahn-Teller (JT) distortions, ferroelectric (FE) distortions and antiferrodistortive (AFD) distortions.

- In the Jahn-Teller distortion the BX_6 octahedron is elongated in one direction and compressed in the two other directions. The Jahn-Teller effect can be attributed to a partial occupations of the two E_g orbitals on the B atom. The E_g orbitals are the $d_{z^2-3r^2}$ orbital and the $d_{x^2-y^2}$ orbitals on the B-atom that point directly towards the oxygen atoms, wn which are antibonding with oxygen. If the axial orbital is filled and the equatorial orbital is empty, the equatorial antibonds are weakened, and contract. Of the occupation is reverted, the axial bonds are contracted.

- In the ferroelectric distortion the B atom in the BX_6 octahedron is displaced from the center. This effect is often attributed to a too small B atom, that is rattling in the octahedral void. This effect is enhanced by large A atoms of the ABO_3 atoms, which expand the lattice and with it the octahedra.[40] For a different point of view see [41].

- in the antiferrodistortive distortion, the nearly rigid BX_6 octahedra are rotated. This effect is often attributed to the fact that the A atom is too small. Tilting allows the a more efficient packing of the octahedra, which improves the attraction of the ions. For a different point of view see [**?**].

  Let us introduce the tolerance factor t defined by Goldschmidt from the bond-distances between the A-Sites (Sr) and the B-sites (Ti) with oxygen.

  $$t \stackrel{\text{def}}{=} \frac{d_{A-O}}{\sqrt{2}d_{B-O}}$$

  If this ratio is unity, the spheres touch in the cubic prototype structure. If the ratio t is smaller than one, the B-O-B bonds are under compressive strain, which is released by octahedral tilting.

12.13.1 Ruddlesden-Popper phases

The perovskite and the rock salt structure can be combined to form superstructures, formed from alternating layers of perovskite and rock salt structure. The structural formula is $A_{n+1}B_nO_{3n+1}$ which corresponds to n double layers of perovskite with composition ABO_3 and one layer of the rock salt structure AO.

12.13.2 Magnéli phases

Magneli phases are layers of rutile and corundum type structures: However, the structure is disputed. (`http://arxiv.org/pdf/cond-mat/0309483`)

M_nO_{2n-1}=M_2O_3+(n-2)VO_2

about oxides see `http://ruby.chemie.uni-freiburg.de/Vorlesung/oxide_0.html`

12.13.3 Intermetallics

Intermetallics are compounds formed from several metal atoms, that, unlike alloys, form ordered structures.

Stichworte

- Laves Phasen
- Heusler Phasen
- Zintl phasen
- Hume Rothery phasen

$L1_0$ is a fcc lattice with alternating planes of A and B atoms. An example is TiAl.

$\vec{T}_1$	$\frac{1}{2}a$	$\frac{1}{2}a$	0
$\vec{T}_2$	$\frac{1}{2}a$	$-\frac{1}{2}a$	0
$\vec{T}_3$	0	0	a
Ti	0	0	0
Al	$\frac{1}{2}a$	0	0

$L1_2$ is an fcc-lattice with one atom type on the cube faces and the other at the cube corners. The parent compound is Ni_3Al.

B2 is a bcc-lattice with one type of atom in the cube center and the other at the cube corners. A compound is NiAl. The B2 structure is essentially the CsCl struycture.

C1 is simply the fluorite structure or the antifluorite structure respectively. Examples are Mg_2Si, Co_2Si, Ni_2Si, Mg_2Ge.

The $D0_3$ structure with composition A_3B can bee seen as a supercell of eight bcc-unit cells, of which the body centers of four cubes are of B-type. examples are Fe_3Al.

$\vec{T}_1$	$2a$	0	0
$\vec{T}_2$	0	$2a$	0
$\vec{T}_3$	0	0	$2a$
Al	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$
Al	$\frac{3}{2}a$	$\frac{3}{2}a$	$\frac{1}{2}a$
Al	$\frac{1}{2}a$	$\frac{3}{2}a$	$\frac{3}{2}a$
Al	$\frac{3}{2}a$	$\frac{1}{2}a$	$\frac{3}{2}a$
Fe	$\frac{1}{2}a$	$\frac{3}{2}a$	$\frac{1}{2}a$
Fe	$\frac{3}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$
Fe	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{3}{2}a$
Fe	$\frac{3}{2}a$	$\frac{3}{2}a$	$\frac{3}{2}a$
Fe	0	0	0
Fe	1	0	0
Fe	0	1	0
Fe	1	1	0
Fe	0	0	1
Fe	1	0	1
Fe	0	1	1
Fe	1	1	1

The A 15 structure with A_3B stoichiometry can be seen as a bcc lattice of Si atoms with Cr-dimers at the faces. The dimers on the faces are oriented such that they make up chains that do not cross each other.

$\vec{T}_1$	a	0	0
$\vec{T}_2$	0	a	0
$\vec{T}_3$	0	0	a
Al	0	0	0
Al	$\frac{1}{2}a$	$\frac{1}{2}a$	$\frac{1}{2}a$
Nb	$\frac{1}{2}a$	$\frac{1}{4}a$	0
Nb	$\frac{1}{2}a$	$\frac{3}{4}a$	0
Nb	0	$\frac{1}{2}a$	$\frac{1}{4}a$
Nb	0	$\frac{1}{2}a$	$\frac{3}{4}a$
Nb	$\frac{1}{4}a$	0	$\frac{1}{2}a$
Nb	$\frac{3}{4}a$	0	$\frac{1}{2}a$

12.14 Additional reading

The article of Wells[42] discusses ordering principles of structures. Interesting is the selection of closed packed structures of type AX*n* with minimum number of neighbors *A*. Perovskite is the simplest ABX_3 structure with closed packing of AX_3 (p464).

Chapter 13

Crystal-Orbital Hamiltonian Populations

One very useful way to analyze local chemical bonding in solids and complex molecules are the projected density of states, Crystal Orbital Overlap (COOP) contributions or Crystal Orbital Hamiltonian Populations.

13.1 Definition of the density of states

These quantities are meant to analyze the band structure energy

$$E_B = \sum_n f_n \epsilon_n$$

which is the sum of orbital energies weighted by their occupations.

The Band structure energy can be rewritten with the help of the total density of states as

$$E_B = \int_{-\infty}^{\infty} d\epsilon\ f(\epsilon-\mu)\epsilon \underbrace{\sum_n \delta(\epsilon-\epsilon_n)}_{D(\epsilon)} = \int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\epsilon f(\epsilon-\mu)$$

The total density of states $D(\epsilon)$ can can now be expressed by a local basisset. Let us consider a non-orthogonal basis set $|\chi_i\rangle$. In order to decompose our wave functions into these local orbitals we introduce projector functions.

$$|\psi_n\rangle = \sum_i |\chi_i\rangle\langle p_i|\psi_n\rangle \tag{13.1}$$

One requirement for this identity to hold is that the basis functions are complete and secondly that

$$\langle p_i|\chi_j\rangle = \delta_{i,j} \tag{13.2}$$

which follows, when we insert the basis function instead of the wave functions in Eq. 13.1

The projector functions can be constructed from a support basis $\langle f_i|$ such from

$$\langle p_i| = \sum_j \bar{Q}_{I,j}\langle f_j| \qquad \text{with} \qquad \sum_k \bar{Q}_{i,k}\langle f_k|\chi_j\rangle = \delta_{i,j}$$

The reader can easily convince himself that the projector functions defined in this way fulfill the biorthogonality condition Eq. 13.2. While nothing forbids to use $\langle f_i| = \langle\chi_i|$, the above formulation is substantially more flexible and more convenient in practice.

The Density of states has the form

$$D(\epsilon)=\sum_n \delta(\epsilon-\epsilon_n)\langle\psi_n|\psi_n\rangle=\sum_{i,j}\underbrace{\sum_n \delta(\epsilon-\epsilon_n)\langle\psi_n|p_i\rangle\langle\chi_i|\chi_j\rangle\langle p_j|\psi_n\rangle}_{D_{j,i}\langle\chi_j|\chi_i\rangle}$$

Similarly we can directly multiply the density of states with the energy

$$D(\epsilon)=\sum_n \delta(\epsilon-\epsilon_n)\langle\psi_n|\psi_n\rangle=\sum_{i,j}\underbrace{\sum_n \delta(\epsilon-\epsilon_n)\langle\psi_n|p_i\rangle\langle\chi_i|\hat{H}|\chi_j\rangle\langle p_j|\psi_n\rangle}_{D_{j,i}\langle\chi_i|\hat{H}|\chi_j\rangle}$$

This defines the density of states in the basis of our basisset as

$$D_{i,j}(\epsilon)=\sum_n \langle p_i|\psi_n\rangle\delta(\epsilon-\epsilon_n)\langle\psi_n|p_j\rangle$$

With the density of states, we obtain the band structure energy as

$$\begin{aligned}E_B&=\sum_{i,j}\int d\epsilon\; f(\epsilon)D_{i,j}(\epsilon)\langle\chi_i|\hat{H}|\chi_j\rangle\\&=\sum_{i,j}\int d\epsilon\; f(\epsilon)D_{i,j}(\epsilon)\epsilon\langle\chi_i|\chi_j\rangle\end{aligned}$$

Thus the density of states yields the complete information about the energy levels, in an energy resolved, spatially resolved and angular momentum resolved manner.

The diagonal elements of $D_{i,j}(\epsilon)\langle\chi_j|\chi_i\rangle$ are called the projected density of states. The off-diagonal elements are called Crystal orbital overlap populations (COOP's). When the Overlap is replaced by the Hamiltonian we obtain the crystal orbital Hamilton population (COHP), $D_{i,j}\langle\chi_j|\hat{H}|\chi_i\rangle$ which directly yields energetic information.

13.2 Example: Formaldehyde

Let us have a look at a very simple example namely formaldehyde H_2co. In Fig. 13.1 the total density of states and COOPs. are shown. The total density of states in this case shows lines for each energy level of the molecule. In order to analyze the CO bond, we determine the COOP between the two mentioned orbitals. Two of the calculated states contribute to bonding. The others are in different symmetry classes. We see that the lower of the two peaks contributes a large negative contribution, indicating a bonding interaction between the two orbitals. That is the two orbitals contribute to this state with the same sign.

The second state has a positive contribution, indicating an anti-bonding interaction. The integral, however shows, that this state does not compensate the bond completely, but about half of the bonding weight is left. This indicates that at higher energy there will be another state with anti-bonding character.

The last two figures show the projection on the p-orbitals of carbon and oxygen respectively. We observe that the bonding state has most of its weight on the carbon atom, while the higher state has almost zero weight on the carbon atom. This state is mostly localized on the oxygen atom.

The results can now be rationalized as follows:

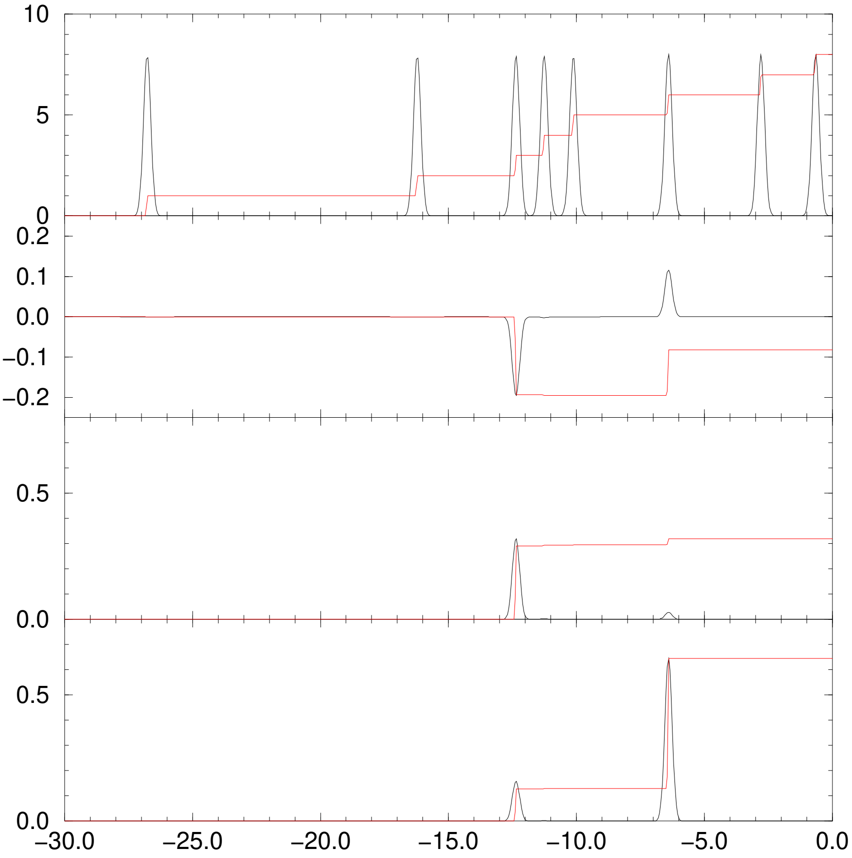


Fig. 13.1: Total density of states (1), COOP between the p orbitals on C and O that lie in the plane of the molecule and that point perpendicular to the CO-axis, and projected density of states respective p-orbital on carbon (3) and oxygen (4) of formaldehyde H_2CO.

13.3 Example: Iron

13.4 Sum rules

If the wave functions, including the unoccupied ones, form a complete set of functions, the expression

$$\sum_n |\psi_n\rangle\langle\psi_n| = \hat{1}$$

If furthermore the basis functions are complete, even not orthonormal, we obtain another representation of the unity operator as

$$\sum_{i,j} |\chi_i\rangle Q_{i,j}\langle\chi_j| = \hat{1}$$

In practice the last equation is only approximative, because the basis functions $|\chi_i\rangle$ do not form a complete basis. In that case one obtains a projection operator on the subspace of the Hilbert space, which is spanned by the basis functions. Nonetheless, if the basis is chosen to capture the main features of the occupied wave functions, the arguments should be sufficient for a qualitative discussion of chemical bonding.

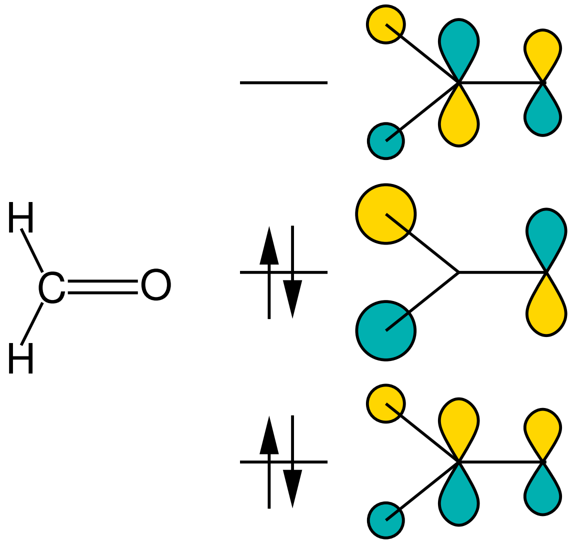


Fig. 13.2: Schematic drawing of the orbitals of formaldehyde discussed in the text.

This allows to write

$$\int_{-\infty}^{\infty} d\epsilon\; D_{j,i}\langle\chi_i|\chi_j\rangle = \sum_{k,l}\underbrace{\langle p_j|\chi_k\rangle}_{\delta_{j,k}} Q_{k,l}\underbrace{\langle\chi_l|p_i\rangle}_{\delta_{l,i}}\langle\chi_i|\chi_j\rangle = Q_{j,i}\langle\chi_i|\chi_j\rangle$$

$$\begin{aligned}
\sum_{k,l}\langle\chi_i|\chi_k\rangle\int_{-\infty}^{\infty} d\epsilon\; D_{k,l}\epsilon\langle\chi_l|\chi_j\rangle &= \sum_{n,k,l}\langle\chi_i|\chi_k\rangle\langle p_j|\psi_n\rangle\langle\psi_n|\hat{H}|\psi_n\rangle\langle\psi_n|p_l\rangle\langle\chi_l|\chi_j\rangle\\
&= \sum_{n,m,k,l}\langle\chi_i|\chi_k\rangle\langle p_j|\psi_n\rangle\langle\psi_n|\hat{H}|\psi_m\rangle\langle\psi_m|p_l\rangle\langle\chi_l|\chi_j\rangle\\
&= \sum_{n,m,o,p,k,l}\langle\chi_i|\chi_k\rangle\langle p_j|\chi_o\rangle Q_{o,p}\langle\chi_p|\hat{H}|\chi_m\rangle Q_{m,n}\langle\chi_n|p_l\rangle\langle\chi_l|\chi_j\rangle\\
&= \sum_{m,p,k,l}\langle\chi_i|\chi_k\rangle Q_{k,p}\langle\chi_p|\hat{H}|\chi_m\rangle Q_{m,l}\langle\chi_l|\chi_j\rangle\\
&= \langle\chi_i|\hat{H}|\chi_j\rangle\\
\int_{-\infty}^{\infty} d\epsilon\; D_{j,i} = \langle p_j|p_i\rangle &= \langle p_j|\chi_k\rangle Q_{k,l}\langle\chi_l|p_i\rangle = Q_{j,i}
\end{aligned}$$

These sum rules allow us to determine the Band structure energy as

$$E_B = \sum_{i,j,k,l}\left[\int_{-\infty}^{\infty} d\epsilon\; f(\epsilon)D_{i,j}(\epsilon)\right]\langle\chi_j|\chi_k\rangle\left[\int_{-\infty}^{\infty} d\epsilon\; D_{k,l}(\epsilon)\epsilon\right]\langle\chi_l|\chi_i\rangle$$

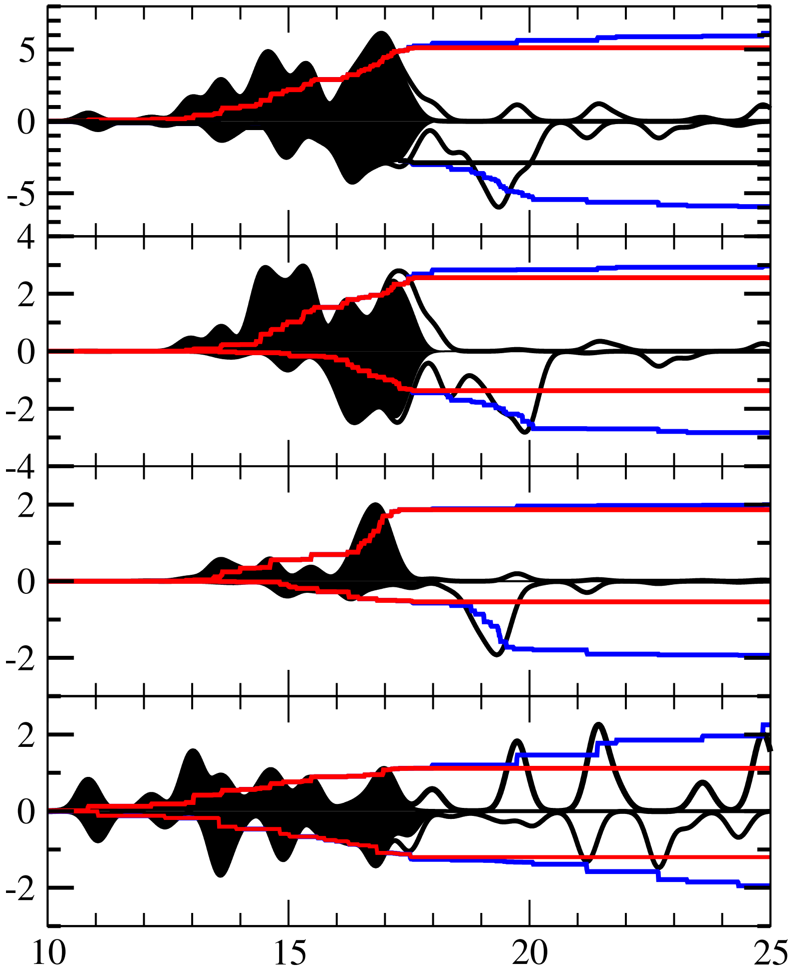


Fig. 13.3: Density of states of bcc-iron: (1) total density of states; (2) density of states on the t2g-d orbitals; (3) density of states on the eg-d orbitals; (4) density of states on the s and p orbitals. Density of states is not to scale. The integrated density of states corresponds to the y-axis. The energy axis is in eV.

Chapter 14

Magnetism

14.1 Quick and Dirty

14.1.1 Driving force for a parallel (ferromagnetic) alignment of spins

The origin of magnetic behavior of materials has, surprisingly, nothing to do with magnetic moments that align due their magnetic fields.

The origin is instead the Pauli principle and the electrostatic interaction between electrons: Electrons with like spin stay apart from each other, while electrons with opposite spin also come close together. Thus, the electrostatic energy due to the Coulomb repulsion between electrons is larger for electrons with opposite spin than for electrons with like spin. Thus, in order to reduce their energy, the electrons tend to align their spins in a parallel orientation.

This a natural tendency of electrons to align with parallel spins is the origin of **Hund's rule**. As electrons fill the shells of an atom, they first orient their spins parallel. Thus the total spins of the atoms is such maximized for each angular momentum shell

Li	Be	B	C	N	O	F	Ne
$\frac{1}{2}\hbar$	0	$\frac{1}{2}\hbar$	$\hbar$	$\frac{3}{2}\hbar$	$\hbar$	$\frac{1}{2}\hbar$	0

Let us now dig a little deeper: Imagine a homogeneous electron gas with a given density. Now consider the density of the remaining electrons as seen by one of the electrons at a specified position. If the material has N electrons our specified electron only sees $N-1$ electrons. Thus it sees the total density with a "hole" around himself. The hole corresponds to exactly the one missing electron.

The Coulomb energy of electrons can be written as

$$E_{e-e}=\frac{1}{2}\int d^3r\int d^3r'\;\frac{e^2n(\vec{r})n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}+\int d^3r\;en(\vec{r})\frac{1}{2}\int d^3r'\;\frac{eh(\vec{r},\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}$$

where $eh(\vec{r},\vec{r'})$ is the positive density of the exchange hole of an electron located at $\vec{r}$. If we compare systems with the same density the first term, called Hartree energy, is identical. Thus, the difference of the Coulomb energy of two systems with the same density can be traced to the attractive interaction of the electrons with their exchange-correlation hole.

If we compare electrostatic energy of a non-magnetic and a ferromagnetic electron gas, the difference is the interaction of the electron with its own exchange-hole. This energy is larger in the ferromagnetic case, because the hole is closer to the specified electron, where the Coulomb interaction is strongest. Thus the electrostatic repulsion between electrons is weaker in the ferromagnetic case.

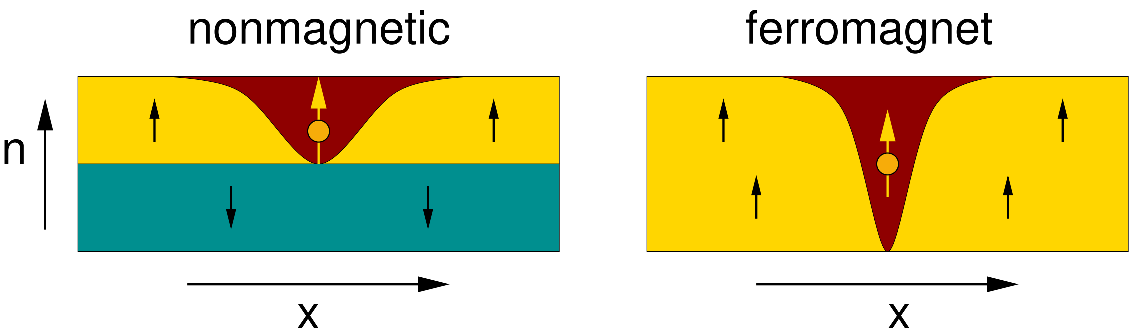


Fig. 14.1: The exchange energy for a magnetic electron gas is favorable compared to a nonmagnetic electron gas. The reason is that the exchange correlation hole of the magnetic electron gas is more compact. The Coulomb attraction of electron and its exchange correlation hole is therefore larger.

14.1.2 Driving force for antiparallel alignment of spins

The story is however not yet complete: Not all materials are magnetic. Even free-electron like metals such as sodium and aluminium are nonmagnetic. The orbitals of molecules are occupied such that electron pair with antiparallel spins, which seems to completely disagree with the above argument.

What is missing is the effect of the kinetic energy. Depending on whether the kinetic energy or the Coulomb energy wins, the material will be nonmagnetic or magnetic.

Let us return to the free electron gas. As we fill the electrons into the orbitals, we note that we need to occupy orbitals with higher energy, if we can only fill one electron into each orbital as in the ferromagnetic case. In the nonmagnetic case we fill the orbitals with an electron pair and less orbitals need to be filled. Thus, from the point of view of one-particle energies, it is always more efficient to fill the lowest orbitals with a singlet pair of two electrons.

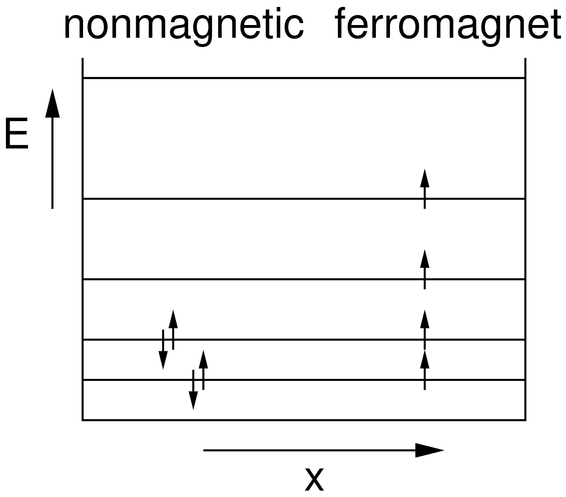
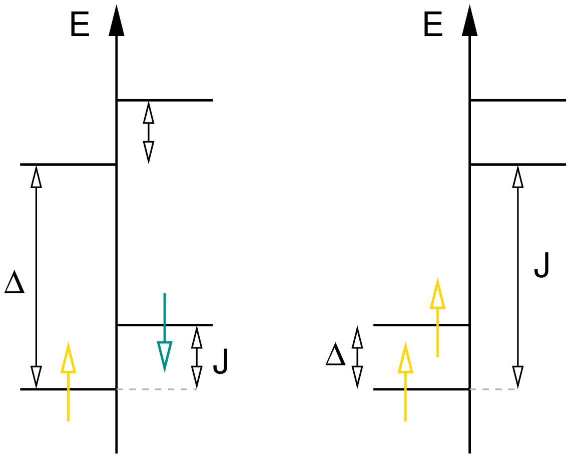


Fig. 14.2: Schematic drawing of the energy levels in a box with non-magnetic and magnetic occupations of the levels. The particle in a box serves as model for a free electron gas.

14.1.3 Taking things together

What did we learn? Magnetism is the result of two antagonistic effects. Coulomb interaction favors parallel spin alignment and kinetic energy favors antiparallel spin alignment. If the energy gap to the next unoccupied state is larger than the gain in Coulomb energy due to parallel spin alignment, the material will become magnetic.



The energy gain due to parallel spin alignment is characterized by the **exchange splitting** J. The loss in kinetic energy can be characterize by the HOMO-LUMO gap Δ in the nonmagnetic case. If the exchange splitting, is larger than the gap Δ the system will favor parallel spin alignment over antiparallel spin alignment.

14.2 Stoner criterion

The **Stoner criterion**[? ?] translates the arguments given above for the balance between exchange splitting and kinetic energy into a simple quantitative criterion for the onset of magnetism in solids.

STONER CRITERION

The **Stoner criterion** says that the system undergoes a magnetic instability if

$$D(\epsilon_F)J > 1 \tag{14.1}$$

that is, if the product of density of states $D(\epsilon_F)$ at the Fermi level and exchange splitting parameter J is larger than one. The density of states is that of the non-magnetic system.

A large density of states at the Fermi level is an indication for a weakly stable system. The large density of states can be lowered by a magnetic transition, which shifts one part of the large density fo states up and the other one down. The exchange splitting parameter J is a measure for a driving force towards a magnetic system. If the latter is sufficiently large, the instability will occur.

While the Stoner criterion is useful to predict whether a material becomes magnetic, it predicts the Curie temperature far too high. Another model to describe the magnetic transition is the Heisenberg model. Whereas the Stoner model assumes a fixed direction of the spins but variations in its size, the Heisenberg model describes interacting spins of fixed size but variable direction. The correct picture is the spin-fluctuation model[43], which considers both longitudinal and transverse fluctuation, that is variations in size and direction. (See also Mohn-Wohlfarth theory[44].)

Exchange energy

We consider the exchange energy as $\Delta E_x = -\frac{1}{2}J(n_\uparrow - n_\downarrow)^2$. The potential is

$$V_\uparrow = \frac{\partial \Delta E_x}{\partial n_\uparrow} = -J(n_\uparrow - n_\downarrow)$$
$$V_\downarrow = \frac{\partial \Delta E_x}{\partial n_\downarrow} = +J(n_\uparrow - n_\downarrow)$$

The potential shift is

$$V_\uparrow - V_\downarrow = 2J\Big(n_\uparrow - n_\downarrow\Big)$$

Kinetic energy

Next, we estimate the kinetic energy: We estimate the work that needs to be done on non-interacting electrons[1] to create a spin density described by $n_\uparrow - n_\downarrow$. The work dW done by a change of a potential is

$$\begin{aligned} dW &= \sum_\sigma n_\sigma dV_\sigma = \frac{1}{2}\Big(n_\uparrow - n_\downarrow\Big)\Big(dV_\uparrow - dV_\downarrow\Big) + \frac{1}{2}\Big(n_\uparrow + n_\downarrow\Big)\underbrace{\Big(dV_\uparrow + dV_\downarrow\Big)}_{\overset{!}{=}0} \\ &= \frac{1}{2}\Big(n_\uparrow - n_\downarrow\Big)\Big(dV_\uparrow - dV_\downarrow\Big) \end{aligned} \qquad (14.2)$$

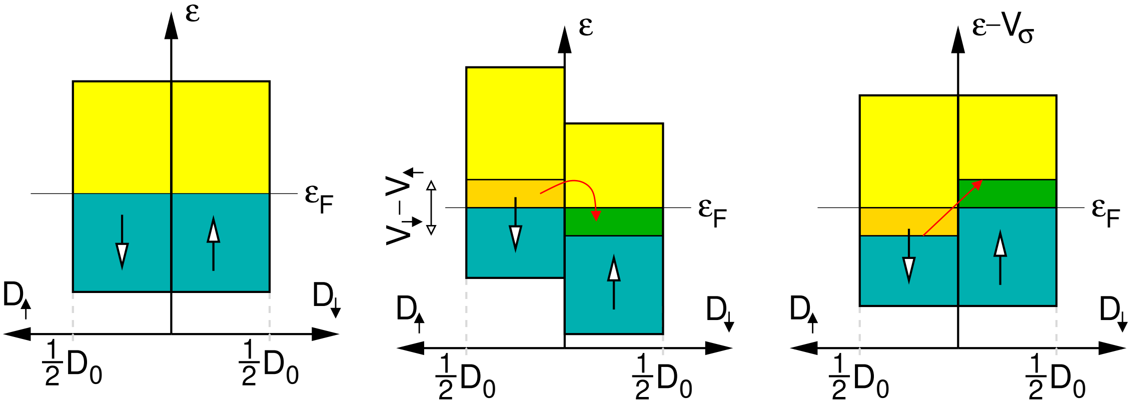


Fig. 14.3: Schematic diagram of the density of states and the increase of the kinetic energy in an applied magnetic field. The spin-down density of states is shown towards the left of the energy axis and the spin-up density of states is shown to the right of the energy axis. Left: The non-magnetic system. Middle: The density of states and occupations for in a magnetic field. Spin up electrons flip their spin in order to take advantage of the lower potential at for down spins. The golden region becomes unoccupied, while the green region becomes occupied. Both spin directions have a common chemical potential μ. Right: the kinetic energy density. If the potential energy contribution to the energy levels removed, we see that the electrons loose kinetic energy.

The spin density produced by the potential is

$$dn_\sigma = D_\sigma(\epsilon_F)dV_\sigma = \frac{1}{2}D_0 dV_\sigma \qquad \Rightarrow \qquad dn_\uparrow - dn_\downarrow = \frac{1}{2}D_0\Big(dV_\uparrow - dV_\downarrow\Big) \qquad (14.3)$$

---

[1] The energy of a non-interacting electron system consists of only kinetic energy and the potential energy in the external potential. If the latter acts in the same way on both spin directions, the energy of the external potential does not change as the spin polarization is created, because this does not affect the total energy. Thus the work done is just the change of the kinetic energy.

Here, we approximated the density of states $D_\sigma(\epsilon)$ by one-half of the total density of states D_0 of the non-spinpolarized system at the Fermi level.

Taken together, we obtain for the work done by polarizing the non-interacting system

$$
\begin{aligned}
dW &= \frac{1}{2}\Big(n_\uparrow - n_\downarrow\Big)\cdot\frac{2}{D_0}\Big(dn_\uparrow - dn_\downarrow\Big)\\
\Rightarrow \qquad \int_0^{n_\uparrow - n_\downarrow} dW &= \frac{1}{2D_0}\Big(n_\uparrow - n_\downarrow\Big)^2
\end{aligned}
\tag{14.4}
$$

Stability of the interacting system

In order to obtain the energy required to polarize the interacting system of electrons, we need to add the exchange splitting.

$$
\int_0^{n_\uparrow - n_\downarrow} dW = \frac{1}{2D_0}\Big(n_\uparrow - n_\downarrow\Big)^2 - \frac{1}{2}J\Big(n_\uparrow - n_\downarrow\Big)^2 = \frac{1}{2}J\Big(n_\uparrow - n_\downarrow\Big)^2\Big(\frac{1}{D_0 J} - 1\Big)
\tag{14.5}
$$

Thus we see that a spin polarization is favorable, if

$$
D_0 J > 1
\tag{14.6}
$$

which is the Stoner criterion Eq. 14.1

14.3 ????

Magnetism is observed most frequently in 3d transition metals and elements with a partially filled f-shell. The atomic orbitals with d- and f-character are strongly bound to the atom core and hardly obverlap. Therefore they form only very flat bands.

Wave functions with d- and f-like angular momentum character are strongly bound to the core

The radial extent of orbitals is closely related to the principal quantum number. The 3-d valence electrons coexist with 4s and 4p valence electrons, but have a radial extent close to the core-like 3s and 3p orbitals. This effect is even stronger for the lanthanides and actinides with a partially filled f-shell: The 4-f electrons in the lanthanides coexist with 6s-, 6p- and 5d- valence electrons. As a consequence the s- and p-electrons either form strong covalent bonds or they even get lost. In a metal they form free electron-like bands with a large band width. This large band width wins over the Coulomb interaction. This is why metals with s- and p-electrons, such as Na dna Al, do not form magnetic metals. Because d and f-electrons are located closer to the nucleus, the form narrow, almost atom-like bands. The loss in kinetic energy for parallel spin-alignment is small. On the other hand the Coulomb interaction is large because the d- and f-electrons are confined into a small spatial region. As a consequence the atomic Hund's rule behavior often survives for transition metals and f-electron materials,leading to magnetic transition metals, and rare earth metals. Consequently lanthanides and actinides are part of the materials used as strong magnets.

14.4 Exchange splitting

The interaction between spins in the Ising model is not of magnetic nature as one might naively expect. The magnetic interaction is long ranged and much smaller than the exchange interactions that play a role in the Ising model.

Let us demonstrate this by exploring the exchange interaction for the minimum Ising model, the Ising dimer, which is also a model for a hydrogen molecule. Consider two sites with one electron on each site. The basis set is built up of a single spatial orbital on each site, ϕ and ψ. Each orbital is a two component spinor, where the upper component corresponds to the spin-up contribution and the lower component is the spin-down contribution so that we can build four one-particle orbitals from two spatial wave functions.

$$
\begin{aligned}
|\uparrow, 0\rangle &= \begin{pmatrix} \phi(r) \\ 0 \end{pmatrix} = \phi(r)\alpha \\
|\downarrow, 0\rangle &= \begin{pmatrix} 0 \\ \phi(r) \end{pmatrix} = \phi(r)\beta \\
|0, \uparrow\rangle &= \begin{pmatrix} \psi(r) \\ 0 \end{pmatrix} = \psi(r)\alpha \\
|0, \downarrow\rangle &= \begin{pmatrix} 0 \\ \psi(r) \end{pmatrix} = \psi(r)\beta
\end{aligned}
$$

where $\alpha = (1, 0)$ and $\beta = (0, 1)$.

Out of the one-particle wave functions, we build two-particle Slater determinants of the type $|\uparrow, \downarrow\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow, 0\rangle \times |0, \downarrow\rangle - |0, \downarrow\rangle \times |\uparrow, 0\rangle\Big)$. Out of the 16 product states that can be build from the four one-particle states, 6 are excluded because they differ only by a sign change from others, and 4 additional ones are excluded because they involve two identical one-particle states, which is prohibited due to Pauli principle. We are left with 6 states:

$$
\begin{aligned}
|\uparrow\downarrow, 0\rangle &= \phi(r_1)\phi(r_2)\frac{\alpha_1\beta_2 - \beta_1\alpha_2}{\sqrt{2}} \\
|\uparrow, \uparrow\rangle, &= \frac{1}{\sqrt{2}}\Big(\phi(r_1)\psi(r_2) - \psi(r_1)\phi(r_2)\Big)\alpha_1\alpha_2 \\
|\uparrow, \downarrow\rangle, &= \frac{1}{\sqrt{2}}\Big(\phi(r_1)\psi(r_2)\alpha_1\beta_2 - \psi(r_1)\phi(r_2)\beta_1\alpha_2\Big) \\
|\downarrow, \uparrow\rangle, &= \frac{1}{\sqrt{2}}\Big(\phi(r_1)\psi(r_2)\beta_1\alpha_2 - \psi(r_1)\phi(r_2)\alpha_1\beta_2\Big) \\
|\downarrow, \downarrow\rangle, &= \frac{1}{\sqrt{2}}\Big(\phi(r_1)\psi(r_2) - \psi(r_1)\phi(r_2)\Big)\beta_1\beta_2 \\
|0, \uparrow\downarrow\rangle &= \psi(r_1)\psi(r_2)\frac{\alpha_1\beta_2 - \beta_1\alpha_2}{\sqrt{2}}
\end{aligned}
$$

Two states, namely $|\uparrow\downarrow, 0\rangle$ and $|0, \uparrow\downarrow\rangle$, correspond to charge transfer states, where one electron has been transferred from one site to another. We exclude them with the argument that the Coulomb repulsion makes these states very unfavorable. This is a model assumption that serves our purpose but is not valid in general.

We evaluate the expectation value of the Hamiltonian.

$$
\begin{aligned}
E = \sum_{\sigma_1,\sigma_2} \int d^3r_1 \int d^3r_2\; \Psi^*(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2) \\
\times \left[\sum_{i=1}^{2}\left(\frac{-\hbar^2}{2m}\nabla_i^2 + V(r_i)\right) + \frac{e^2}{4\pi\epsilon_0}\frac{1}{|r_1 - r_2|}\right]\psi(\vec{r}_1, \sigma, \vec{r}_2, \sigma)
\end{aligned}
\tag{14.7}
$$

Let us now consider only Slater determinants, which implies that we make the so-called **Hartree-Fock approximation**

$$
\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2) = \frac{1}{\sqrt{2}}\Big(\chi_a(\vec{r}_1, \sigma_1)\chi_b(\vec{r}_2, \sigma_2) - \chi_b(\vec{r}_1, \sigma_1)\chi_a(\vec{r}_2, \sigma_2)\Big) \tag{14.8}
$$

where χ_a and χ_b are orthonormal spin orbitals. With Slater determinants the total energy has the form.

$$E = \sum_{j\in\{a,b\}} \underbrace{\left[\sum_{\sigma}\int d^3r\; \chi_j(\vec{r},\sigma)\left(\frac{-\hbar^2}{2m}\nabla_i^2 + V(r_i)\right)\chi_j(\vec{r},\sigma)\right]}_{h_{a,b}} \tag{14.9}$$

$$\sum_{\sigma_1,\sigma_2}\int d^3r_1 \int d^3r_2\; \Psi^*(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma_2)$$
$$\times\left[\sum_{i=1}^{2}\left(\frac{-\hbar^2}{2m}\nabla_i^2 + V(r_i)\right) + \frac{e^2}{4\pi\epsilon_0}\frac{1}{|r_1-r_2|}\right]\psi(\vec{r}_1,\sigma,\vec{r}_2,\sigma) \tag{14.10}$$

Note here, that the operator yielding the charge density is proportional to the unity operator in spin space.

$$\begin{aligned}H = &\sum_i\left(\frac{-\hbar^2}{2m}\nabla_i^2 + V(r_i)\right)\begin{pmatrix}1\,0\\0\,1\end{pmatrix}\\ &+\frac{e^2}{4\pi\epsilon}\frac{1}{|r_1-r_2|}\begin{pmatrix}|r_1\rangle\langle r_1| & 0\\ 0 & |r_1\rangle\langle r_1|\end{pmatrix}\begin{pmatrix}|r_2\rangle\langle r_2| & 0\\ 0 & |r_2\rangle\langle r_2|\end{pmatrix}\end{aligned}$$

we need four variables

$$\begin{aligned}\epsilon &= \int d^3r\; \phi_i^*(r)\Big[\frac{-\hbar^2}{2m}\nabla^2 + V(r)\Big]\phi_i(r)\\ U &= \int dr\int dr'\; \frac{\phi_1^*(r)\phi(r)\phi_1^*(r)\phi(r)}{|r-r'|}\\ K &= \int dr\int dr'\; \frac{\phi_1^*(r)\phi(r)\phi_2^*(r)\psi(r)}{|r-r'|}\\ J &= \int dr\int dr'\; \frac{\phi_1^*(r)\psi(r)\phi_2^*(r)\phi(r)}{|r-r'|}\end{aligned}$$

In order to diagonalize the Hamiltonian, we form angular momentum eigenstates

$$\begin{aligned}|\ell=0,m=0\rangle &= \frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle - |\downarrow\uparrow\rangle\Big)\\ |\ell=1,m=1\rangle &= |\uparrow,\uparrow\rangle\\ |\ell=1,m=0\rangle &= \frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle + |\downarrow\uparrow\rangle\Big)\\ |\ell=1,m=-1\rangle &= |\downarrow,\downarrow\rangle\end{aligned}$$

The energies of the eigenstates can then be evaluated as the diagonal elements of the Hamiltonian

$$\begin{pmatrix}\langle\uparrow\uparrow|\\ \langle\uparrow\downarrow|\\ \langle\downarrow\uparrow|\\ \langle\downarrow\downarrow|\end{pmatrix} H \begin{pmatrix}|\uparrow\uparrow\rangle\\ |\uparrow\downarrow\rangle\\ |\downarrow\uparrow\rangle\\ |\downarrow\downarrow\rangle\end{pmatrix} = \begin{pmatrix}(K-J), & 0, & 0, & 0\\ 0, & K, & -J, & 0\\ 0, & -J, & K, & 0\\ 0, & 0, & 0, & (K-J)\end{pmatrix}$$

We obtain the eigenstates as

$$\begin{aligned}H|\uparrow,\uparrow\rangle &= (K-J)|\uparrow,\uparrow\rangle\\ H\frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle\Big) &= (K-J)\frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle\Big)\\ H\frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle\Big) &= (K+J)\frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle\Big)\\ H|\downarrow,\downarrow\rangle &= K-J|\downarrow,\downarrow\rangle\end{aligned}$$

The **Coulomb integral** K appears on all diagonal elements, because all states have equal charge density, and only spins are interchanged.

The two states

$$\begin{aligned}\frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle-|\downarrow,\uparrow\rangle\Big)&=\frac{1}{2}\Big(\phi(r_1)\psi(r_2)+\psi(r_1)\phi(r_2)\Big)\Big(\alpha_1\beta_2-\beta_1\alpha_2\Big)\\ \frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle+|\downarrow,\uparrow\rangle\Big)&=\frac{1}{2}\Big(\phi(r_1)\psi(r_2)-\psi(r_1)\phi(r_2)\Big)\Big(\alpha_1\beta_2+\beta_1\alpha_2\Big)\end{aligned}$$

are the singlet state with $S_{z,1}+S_{z,2}=0$ and the triplet state $S_{z,1}+S_{z,2}=\hbar$.

Thus, I can write the Hamiltonian as

$$H=K-J\sigma_1\sigma_2$$

The coupling among the spins is the **exchange coupling** J. This Hamiltonian shows clearly that two electrons will prefer to be parallel to each other, if the exchange coupling is positive.

14.5 Zener's theory

(Zener described that the coupling between d-shells is normally antiferromagnetic and that it becomes ferromagnetic by coupling to conduction electrons. [45] The Arguments may be worth to mention.)

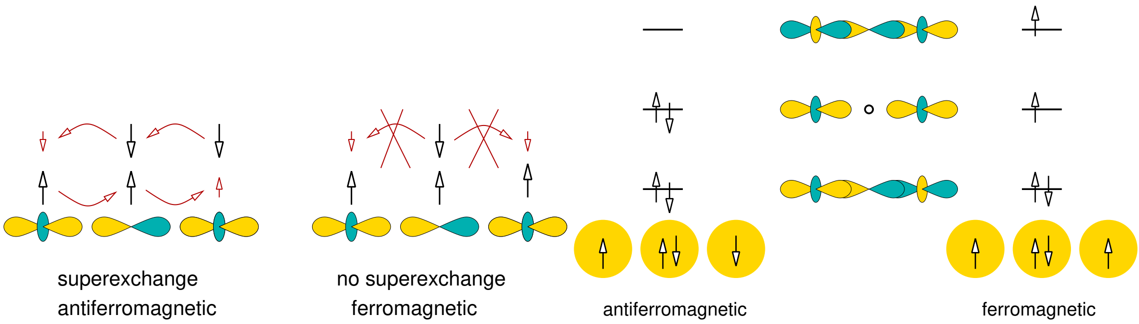
14.6 Super-exchange

Super-exchange was introduced by Kramers[46]. A theory has been worked out by P.W. Anderson[47]

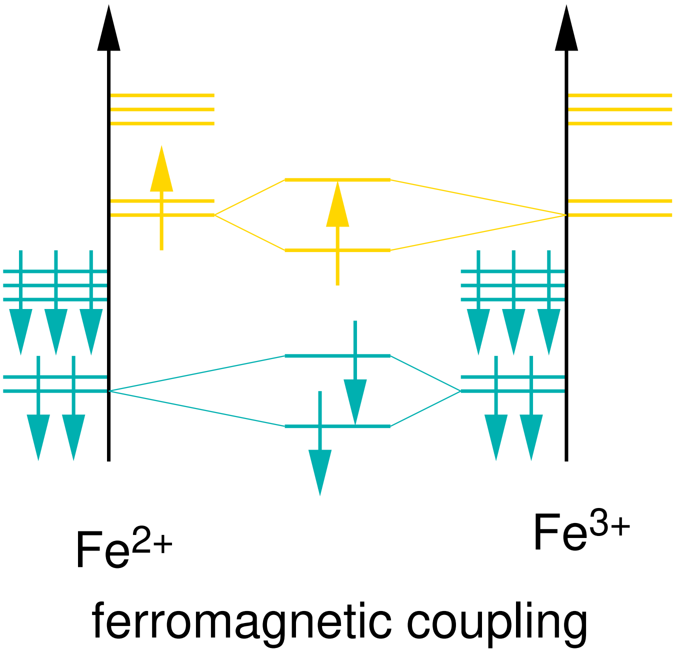
Consider two atoms, which have a spin, which is connected by a non-metal. This is a bonding scheme that is found in most transition metal oxides and also in iron-sulfur complexes, that play an important role in biochemistry.

The spin-alignment of the two magnetic atoms is typically antiferromagnetic.

The simple picture is that an electron on one d-site can hop via the intermediate atom to the other side. According to Heisenberg's uncertainty principle, this lowers the kinetic energy. However, hopping to the other side is only possible if the orbital, to which the electrons is going to hop, is empty, which is the case if the ions are antiferromagnetically coupled. If the spins are ferromagnetically coupled, all three orbitals of one spin direction are occupied, so that these cannot hop to neighboring states. Only the minority electron on the bridging atom could hop to the transition metal sites. This however would induce are charge transfer from the anion to the cation, which is too costly in energy.



If the orbital on the right side would be empty, so that only three electrons are in the system, the majority electrons could migrate to the other side anyway, and ferromagnetic coupling would result.



If the orbital on the right side would not couple to the bridging atom, because it is orthogonal to it, the super-exchange mechanism would not work, because the electron could not hop to the other side, because there is no overlap matrix element.

One of the original papers on the super-exchange mechanism has been written by PW Anderson[47].

14.7 Double exchange

The term **double exchange** has been coined by Zener[48] in 1951 to explain the relation of ferromagnetism and electron conductivity in manganites.

The intention of the model is to describe the transport in $Ca_xLa_{1-x}MnO_3$. This material is a perovskite with x Mn^{4+} ions and $1-x$ Mn^{3+} ions. For the parent compound with $x=1$, i.e. $Ca^{2+}Mn^{4+}O_3^{2-}$ the Mn ion has 3 valence electrons left, which fill the non-bonding t_{2g} d-orbitals of the majority spin direction. $CaMnO_3$ is an antiferromagnetic insulator.

When $CaMnO_3$ is doped by La, additional electrons are introduced onto the Mn-d-shell and Mn^{3+} are formed. The "conduction band" absorbing the additional electrons of the Mn^{3+} ions is made of the antibonding e_g orbitals of the majority spin direction. The e_g orbitals are the ones that point directly towards the oxygen neighbors. While $CaMnO_3$ is an antiferromagnetic insulator, the doped material is ferromagnetic and conducting.

The double-exchange mechanism describes how two metal atoms, that are connected only by an intermediate closed-shell anion, are ferromagnetically coupled and how they obtain metallic behavior. The underlying model is that of two transition-metal orbitals connected by a central orbital. The central atom is filled and there is a single electron in the transition metal orbitals.

The term double exchange describes the delocalization of the electron over the two sites by an indirect mechanism. Namely first an electron is kicked out from the central site to the empty d-orbital, forming a virtual orbital, and secondly the electron from the previously filled d-orbital drops into the central site. This double exchange mechanism only works if the spin on the two d-sites have the same spin. Thus this mechanism stabilizes ferromagnetism.

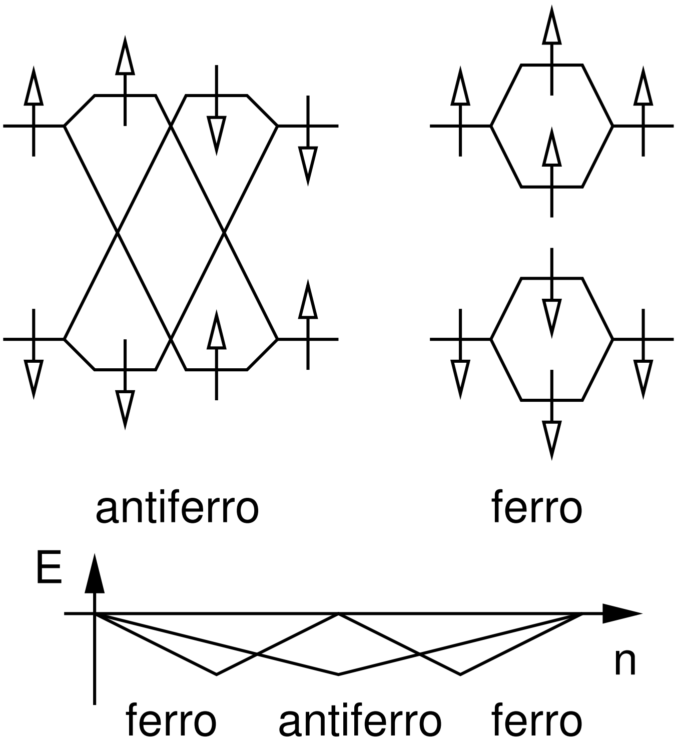


Fig. 14.4: Scheme to demonstrate the competition between ferromagnetism and antiferromagnetism. If there are two electrons on the sytem they can be accomodated in bonding orbitals only when the two atoms are antiferromagnetic. No covalen binding survives if the two atoms are ferromagnetically aligned. If, on the other hand, there is only one orbital in the bonding or the antibonding orbital, that is $n = 1$ and $n = 3$, the ferromagnetic coupling is more favorate, because the hybridization is stronger if the atomic levels are closer together. Bottom: schematic binding energy as function of number of electrons.

The ground state will be a superposition of the four relevant three-particle states

$$
\begin{aligned}
|\chi_1\rangle &= \hat{d}^\dagger_{1,\uparrow}\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle \\
|\chi_2\rangle &= \hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle \\
|\chi_3\rangle &= \hat{d}^\dagger_{1,\downarrow}\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle \\
|\chi_4\rangle &= \hat{d}^\dagger_{1,\downarrow}\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle
\end{aligned}
$$

The operators $\hat{c}^\dagger_\sigma$ create an electron on the central oxygen atom, while $d^\dagger_{1,\sigma}$ creates an electron on the left transition metal atom, while $d^\dagger_{2,\sigma}$ creates an electron on the right transition metal atom.

Because the Hamiltonian has rotational symmetry regarding the spin, the Hamiltonian does not couple states with different total spin. That is the Hamiltonian is block diagonal with one block describing states with $S = +\frac{\hbar}{2}$ and the other block containing states with $S = -\frac{\hbar}{2}$. The problem furthermore has parity symmetry, which blocks the states into mirror symmetric and antisymmetric states. Thus the eigenstates will have the form

$$
\begin{aligned}
|\psi_1\rangle &= \frac{1}{\sqrt{2}}\left(\hat{d}^\dagger_{1,\uparrow} + \hat{d}^\dagger_{2,\uparrow}\right)\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle \\
|\psi_2\rangle &= \frac{1}{\sqrt{2}}\left(\hat{d}^\dagger_{1,\uparrow} - \hat{d}^\dagger_{2,\uparrow}\right)\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle \\
|\psi_3\rangle &= \frac{1}{\sqrt{2}}\left(\hat{d}^\dagger_{1,\downarrow} + \hat{d}^\dagger_{2,\downarrow}\right)\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle \\
|\psi_4\rangle &= \frac{1}{\sqrt{2}}\left(\hat{d}^\dagger_{1,\downarrow} - \hat{d}^\dagger_{2,\downarrow}\right)\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle
\end{aligned}
$$

However, there are no direct Hamilton matrix elements among the basis states $|\chi_1\rangle, \ldots, |\chi_4\rangle$. The presence of a coupling term is important to drive delocalization of the d-electron between the two sites. An effective coupling, however, comes about by including virtual states, for which one of the electrons is kicked out from the central atom and placed into an empty d-orbital.

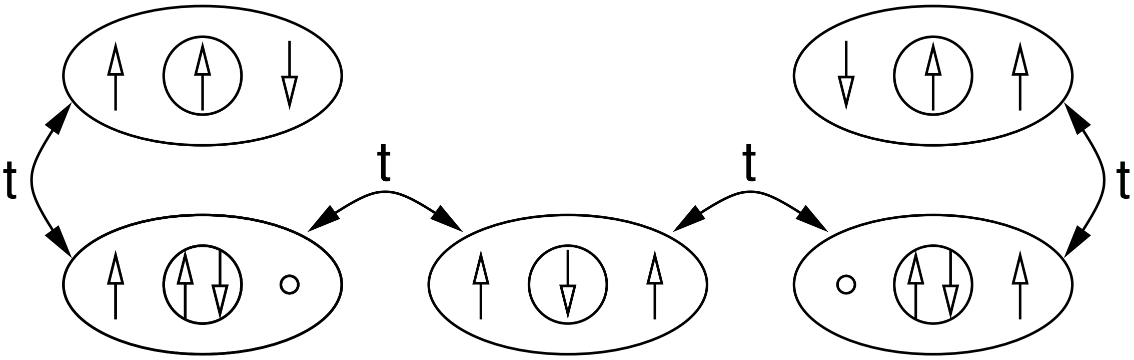


Fig. 14.5: Scheme describing the Slater determinants obtained by hopping from one of the real states. We have excluded states with double occupation on a peripheral site. This neglect has only quantitative consequences. Note that the hopping into an antiferromagnetic virtual orbital does not allow a coupling between two real configuration. It only renormalizes the energy.

Consider a Hamiltonian between two real and a virtual orbital

$$
\hat{H} = \begin{pmatrix} \hat{d}^\dagger_{1,\uparrow}\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle \\ \hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle \\ \hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle \end{pmatrix}
\begin{pmatrix} 2\epsilon_0+\epsilon_d & 0 & t \\ 0 & 2\epsilon_0+\epsilon_d & t \\ t & t & \epsilon_0+2\epsilon_d \end{pmatrix}
\begin{pmatrix} \langle\mathcal{O}|\hat{c}_\downarrow\hat{c}_\uparrow\hat{d}_{1,\uparrow} \\ \langle\mathcal{O}|\hat{c}_\downarrow\hat{c}_\uparrow\hat{d}_{2,\uparrow} \\ \langle\mathcal{O}|\hat{c}_\downarrow\hat{d}_{2,\uparrow}\hat{d}_{1,\uparrow} \end{pmatrix}
$$

This is the same Hamiltonian as that for the three-center bond. The difference is that here it is a many-particle Hamiltonian. The ground state can therefore be written easily as

$$
|\Psi\rangle = \frac{1}{\sqrt{2}}\left(\hat{d}^\dagger_1 + \hat{d}^\dagger_2\right)\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle\cos(\alpha) + \hat{d}^\dagger_{1\uparrow}\hat{d}^\dagger_{2\uparrow}\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle\sin(\alpha)
$$

where α is a small parameter that is still to be determined.

We can treat the third Slater determinant as virtual orbital, and down fold it.

$$
\begin{aligned}
0 = (\boldsymbol{H} - \epsilon \boldsymbol{1})\, \vec{c} &= \begin{pmatrix} \boldsymbol{H}_0 - \epsilon \boldsymbol{1} & \boldsymbol{t} \\ \boldsymbol{t}^\dagger & \epsilon_v - \epsilon \end{pmatrix} \begin{pmatrix} \vec{c}_0 \\ c_v \end{pmatrix} \\
c_v &= -\frac{1}{\epsilon_v - \epsilon} \boldsymbol{t}^\dagger \vec{c}_0 \\
0 &= \left[\boldsymbol{H}_0 - \epsilon \boldsymbol{1} - \boldsymbol{t} \frac{1}{\epsilon_v - \epsilon} \boldsymbol{t}^\dagger \right] \vec{c}_0 \\
H_{eff}(\epsilon) &= \begin{pmatrix} 2\epsilon_0 + \epsilon_d - \frac{t^2}{\epsilon_v - \epsilon} & -\frac{t^2}{\epsilon_v - \epsilon} \\ -\frac{t^2}{\epsilon_v - \epsilon} & 2\epsilon_0 + \epsilon_d - \frac{t^2}{\epsilon_v - \epsilon} \end{pmatrix} \\
\epsilon &\approx 2\epsilon_0 + \epsilon_d - \frac{t^2}{\epsilon_d - \epsilon_0} \pm \frac{t^2}{\epsilon_d - \epsilon_0}
\end{aligned}
$$

In a one-particle picture the Hamiltonian would describe a three-center bond, which a singly occupied non-bonding orbital.

The energy splitting between the groundstate and the first excited state is a measure for the frequency with which an electron oscillates from one to the other site.

$$
\hbar\omega = \frac{2t^2}{\epsilon_d - \epsilon_0} = \frac{2\pi\hbar}{T}
$$

Thus hopping frequency is

$$
\frac{1}{T} = \frac{2t^2}{2\pi\hbar(\epsilon_d - \epsilon_0)}
$$

with the distance L of the two peripheral sites, we can extract a diffusion constant from

$$
D = \frac{1}{6} \frac{L^2}{T} = \frac{t^2 L}{6\pi\hbar(\epsilon_d - \epsilon_0)}
$$

Using the Einstein relation

$$
\sigma = \frac{1}{k_B T} n q^2 D
$$

we obtain an estimate for the electrical conductivity σ or a given number density n of carriers.

The Zener double exchange does not require any electron-electron interaction, but only a hopping from one d-orbital to the other via an occupied central orbital.

Explicit derivation

This process must allow for an effective hopping from one metal ion to another. This process proceeds by first knocking out an electron from the anion into the second metal atom. If the electron has the same spin as the previous metal electron, the latter can occupy the level on the anion, that just lost its electron. As a result the electron hopped from one site to the other in a process where actually two electrons where exchanged between two pairs of orbitals.

The Hamiltonian is

$$
\hat{H} = \sum_{i,\sigma} \epsilon_d \hat{d}^\dagger_{i,\sigma} \hat{d}_{i,\sigma} + \sum_{\sigma} \epsilon_0 \hat{c}^\dagger_{i,\sigma} \hat{c}_{i,\sigma} + \sum_{i,\sigma} t \left(\hat{d}^\dagger_{i,\sigma} \hat{c}_\sigma + \hat{c}^\dagger_\sigma \hat{d}_{i,\sigma} \right) + \sum_i (U - J) \hat{d}^\dagger_{i,\uparrow} \hat{d}_{i,\uparrow} \hat{d}^\dagger_{i,\downarrow} \hat{d}_{i,\downarrow}
$$

$$
\begin{aligned}
\hat{H}|\chi_1\rangle &= |\chi_1\rangle\,(\epsilon_d + 2\epsilon_0) + \Big(-\hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{1,\downarrow}\hat{c}^\dagger_{\uparrow}|0\rangle - \hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_{\downarrow}|0\rangle - \hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\uparrow}|0\rangle\Big)\,t\\
\hat{H}|\chi_2\rangle &= |\chi_2\rangle\,(\epsilon_d + 2\epsilon_0) + \Big(-\hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_{\downarrow}|0\rangle + \hat{d}^\dagger_{1,\downarrow}\hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_{\uparrow}|0\rangle + \hat{d}^\dagger_{2,\uparrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\downarrow}|0\rangle\Big)\,t\\
\hat{H}|\chi_3\rangle &= |\chi_3\rangle\,(\epsilon_d + 2\epsilon_0) + \Big(-\hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{1,\downarrow}\hat{c}^\dagger_{\downarrow}|0\rangle + \hat{d}^\dagger_{1,\downarrow}\hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_{\downarrow}|0\rangle - \hat{d}^\dagger_{1,\downarrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\uparrow}|0\rangle\Big)\,t\\
\hat{H}|\chi_4\rangle &= |\chi_1\rangle\,(\epsilon_d + 2\epsilon_0) + \Big(-\hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\downarrow}|0\rangle + \hat{d}^\dagger_{1,\downarrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\uparrow}|0\rangle - \hat{d}^\dagger_{2,\uparrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\downarrow}|0\rangle\Big)\,t
\end{aligned}
$$

We see that the four states are degenerate, that is $\langle\chi_i|\hat{H}|\chi_i\rangle$ is the same for all orbitals. However, these states are no eigenstates, but the Hamiltonian scatters the electrons out of the central atom into the neighbors. The scattering however does not change the spin. The Hamiltonian is block diagonal in the states with different total spin.

Thus we need only consider two out of the four states at a time. We begin with $|\chi_1\rangle$ and $|\chi_2\rangle$. In order to obtain eigenstates we need to include the states with total ↑ spin and two electrons in the d-orbitals and one in the central atom

$$
\begin{aligned}
|\phi_1\rangle &= \hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{1,\downarrow}\hat{c}^\dagger_{\uparrow}|0\rangle\\
|\phi_2\rangle &= \hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_{\downarrow}|0\rangle\\
|\phi_3\rangle &= \hat{d}^\dagger_{1,\uparrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\uparrow}|0\rangle\\
|\phi_4\rangle &= \hat{d}^\dagger_{1,\downarrow}\hat{d}^\dagger_{2,\uparrow}\hat{c}^\dagger_{\uparrow}|0\rangle\\
|\phi_5\rangle &= \hat{d}^\dagger_{2,\uparrow}\hat{d}^\dagger_{2,\downarrow}\hat{c}^\dagger_{\uparrow}|0\rangle
\end{aligned}
$$

There are five orbitals. In two of the states $|\phi_1\rangle, |\phi_5\rangle$ there are two electrons on one d-atom. In the remaining states there is one electron of each neighbor. Two of those states $|\phi_3\rangle, |\phi_4\rangle$ describe antiferromagnetically coupled electrons, and one, $|\phi_2\rangle$ describes ferromagnetic coupling.

$$
\begin{aligned}
\hat{H}|\chi_1\rangle &= |\chi_1\rangle\,(\epsilon_d + 2\epsilon_0) + (-|\phi_1\rangle - |\phi_2\rangle - |\phi_3\rangle)\,t\\
\hat{H}|\chi_2\rangle &= |\chi_2\rangle\,(\epsilon_d + 2\epsilon_0) + (-|\phi_2\rangle + |\phi_4\rangle + |\phi_5\rangle)\,t
\end{aligned}
$$

The energies of the new states are

$$
\begin{aligned}
\langle\phi_1|\hat{H}|\phi_1\rangle &= 2\epsilon_d + \epsilon_0 + U - J\\
\langle\phi_2|\hat{H}|\phi_2\rangle &= 2\epsilon_d + \epsilon_0\\
\langle\phi_3|\hat{H}|\phi_2\rangle &= 2\epsilon_d + \epsilon_0\\
\langle\phi_4|\hat{H}|\phi_2\rangle &= 2\epsilon_d + \epsilon_0\\
\langle\phi_5|\hat{H}|\phi_2\rangle &= 2\epsilon_d + \epsilon_0 + U - J
\end{aligned}
$$

Next we downfold the states $|\phi_i\rangle$:

$$
\begin{aligned}
\begin{pmatrix} \boldsymbol{H}_\chi - \epsilon\boldsymbol{1} & \boldsymbol{Q}\\ \boldsymbol{Q}^\dagger & \boldsymbol{H}_\phi - \epsilon\boldsymbol{1}\end{pmatrix}\begin{pmatrix}\vec{c}_\chi\\ \vec{c}_\phi\end{pmatrix} &= 0\\
\boldsymbol{Q}^\dagger\vec{c}_\chi + (\boldsymbol{H}_\phi - \epsilon\boldsymbol{1})\vec{c}_\phi &= 0\\
\vec{c}_\phi &= -(\boldsymbol{H}_\phi - \epsilon\boldsymbol{1})^{-1}\boldsymbol{Q}^\dagger\vec{c}_\chi\\
\left[\boldsymbol{H}_\chi - \epsilon\boldsymbol{1} - \boldsymbol{Q}(\boldsymbol{H}_\phi - \epsilon\boldsymbol{1})^{-1}\boldsymbol{Q}^\dagger\right]\vec{c}_\chi &= 0
\end{aligned}
$$

The states $|\phi_i\rangle$ act like an additional interaction between the states $|\chi_i\rangle$.

$$
\begin{aligned}
&-\boldsymbol{Q}(\boldsymbol{H}_\phi-\epsilon\boldsymbol{1})^{-1}\boldsymbol{Q}^\dagger\\
=&-\begin{pmatrix}-t & -t & -t & 0 & 0\\ 0 & -t & 0 & t & t\end{pmatrix}\\
&\cdot\begin{pmatrix}2\epsilon_d+\epsilon_0+U-\epsilon & 0 & 0 & 0 & 0\\ 0 & 2\epsilon_d+\epsilon_0-\epsilon & 0 & 0 & 0\\ 0 & 0 & 2\epsilon_d+\epsilon_0-\epsilon & 0 & 0\\ 0 & 0 & 0 & 2\epsilon_d+\epsilon_0-\epsilon & 0\\ 0 & 0 & 0 & 0 & 2\epsilon_d+\epsilon_0+U-\epsilon\end{pmatrix}^{-1}\begin{pmatrix}-t & 0\\ -t & -t\\ -t & 0\\ 0 & t\\ 0 & t\end{pmatrix}\\
=&-\begin{pmatrix}-t & -t & -t & 0 & 0\\ 0 & -t & 0 & t & t\end{pmatrix}\begin{pmatrix}\frac{-t}{2\epsilon_d+\epsilon_0+U-\epsilon} & 0\\ \frac{-t}{2\epsilon_d+\epsilon_0-\epsilon} & \frac{-t}{2\epsilon_d+\epsilon_0-\epsilon}\\ \frac{-t}{2\epsilon_d+\epsilon_0-\epsilon} & 0\\ 0 & \frac{t}{2\epsilon_d+\epsilon_0-\epsilon}\\ 0 & \frac{t}{2\epsilon_d+\epsilon_0+U-\epsilon}\end{pmatrix}^{-1}\\
=&-\begin{pmatrix}\frac{t^2}{2\epsilon_d+\epsilon_0+U-\epsilon}+\frac{2t^2}{2\epsilon_d+\epsilon_0-\epsilon} & \frac{t^2}{2\epsilon_d+\epsilon_0-\epsilon}\\ \frac{t^2}{2\epsilon_d+\epsilon_0-\epsilon} & \frac{t^2}{2\epsilon_d+\epsilon_0+U-\epsilon}+\frac{2t^2}{2\epsilon_d+\epsilon_0-\epsilon}\end{pmatrix}
\end{aligned}
$$

If we insert $\epsilon\approx 2\epsilon_0+\epsilon_d$ the additional interaction has the form

$$
=-\begin{pmatrix}\frac{t^2}{\epsilon_d+U-\epsilon_0}+\frac{2t^2}{\epsilon_d-\epsilon_0} & \frac{t^2}{\epsilon_d-\epsilon_0}\\ \frac{t^2}{\epsilon_d-\epsilon_0} & \frac{t^2}{\epsilon_d+U-\epsilon_0}+\frac{2t^2}{\epsilon_d-\epsilon_0}\end{pmatrix}
$$

Thus the ground state has the form

$$
\frac{1}{\sqrt{2}}\left(|\chi_1+\chi_2\rangle\right)=\left(\hat{d}^\dagger_{1,\uparrow}+\hat{d}^\dagger_{2,\uparrow}\right)\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle
$$

Thus we obtain an effective ferromagnetic interaction of the two peripheral sites and an electron that is delocalized over both sites.

The first excited state has the form

$$
\frac{1}{\sqrt{2}}\left(|\chi_1\rangle-|\chi_2\rangle\right)=\left(\hat{d}^\dagger_{1,\uparrow}-\hat{d}^\dagger_{2,\uparrow}\right)\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|0\rangle
$$

The excitation energy is $2t^2/(\epsilon_d-\epsilon_0)$. Thus the electron oscillates between the two sites with frequency $\hbar\omega=2t^2/(\epsilon_d-\epsilon_0)$. If the level separation is sufficiently low, that frequency will be large, which is indicative of a metallic behavior. From the frequency and the distance of the neighboring sites one can even estimate the diffusion constant of the electron in a lattice and thus the conductivity.

Appendices

Appendix A

Polyhedra in paper

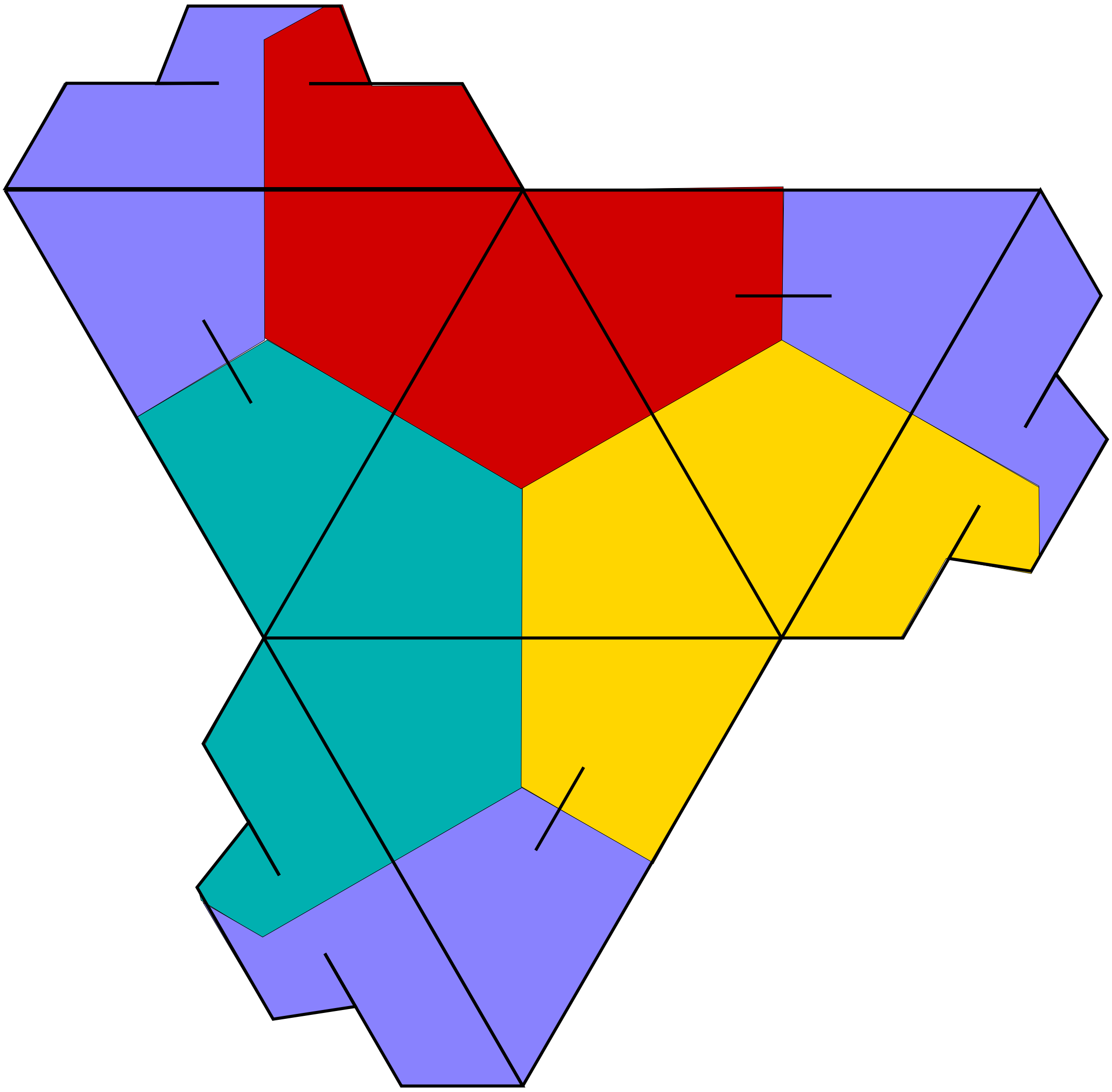


Fig. A.1: Pattern for constructing a tetrahedron

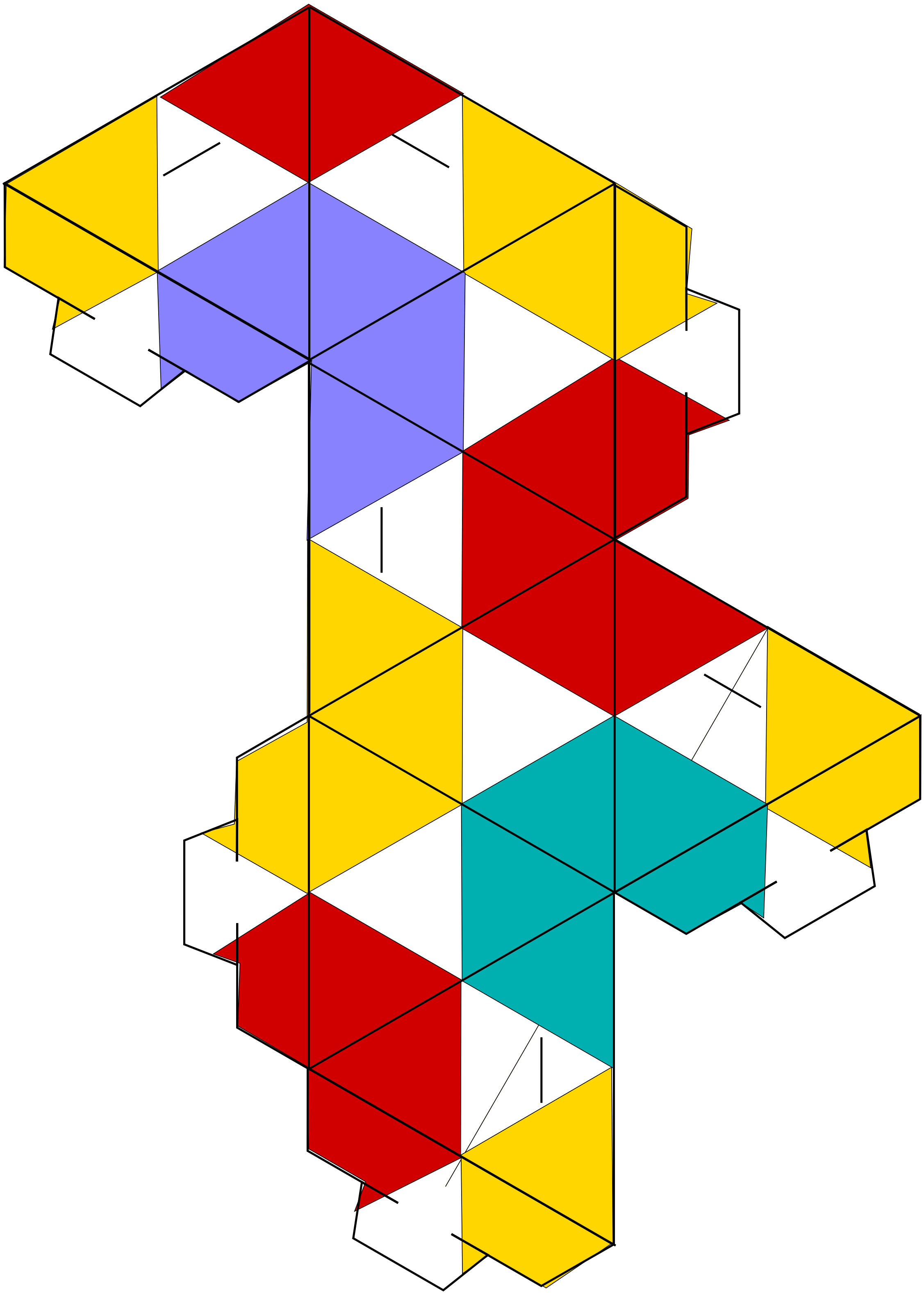


Fig. A.2: Pattern for constructing an octahedron

Appendix B

Justification of the Wolfsberg formula

Using the empirical Wolfsberg formula Eq. 2.6

$$t = k\frac{\bar{\epsilon}_1 + \bar{\epsilon}_2}{2}\Delta$$

we can map the Schrödinger equation with a Hamiltonian and Overlap matrix given by Eq. 2.4

$$\mathbf{H} = \begin{pmatrix} \bar{\epsilon}_1 & t \\ t & \bar{\epsilon}_2 \end{pmatrix} \qquad \text{and} \qquad \mathbf{O} = \begin{pmatrix} 1 & \Delta \\ \Delta & 1 \end{pmatrix} \tag{B.1}$$

to first order in Δ onto one with orthogonal Basis functions.

B.1 Approximate orthonormalization

The overlap operator can be made equal to unity by a transformation of the basis orbitals

$$\begin{pmatrix} |\chi'_1\rangle \\ |\chi'_2\rangle \end{pmatrix} = \begin{pmatrix} |\chi_1\rangle \\ |\chi_2\rangle \end{pmatrix} \mathbf{O}^{\frac{1}{2}}$$

Let us first form the square root of the overlap matrix to first order in Δ. We start with an ansatz for the overlap matrix

$$\mathbf{O}^{\frac{1}{2}} = 1 + \Delta\mathbf{A} + O(\Delta^2)$$

which defines the matrix $\mathbf{A}$.

The matrix $\mathbf{A}$ has to obey the equation

$$\mathbf{O} = (\mathbf{1} + \Delta\mathbf{A})(\mathbf{1} + \Delta\mathbf{A}) + O(\Delta^2) = \mathbf{1} + 2\Delta\mathbf{A} + O(\Delta^2)$$
$$\Rightarrow \mathbf{1} + \Delta\mathbf{A} = \begin{pmatrix} 1 & \frac{1}{2}\Delta \\ \frac{1}{2}\Delta & 1 \end{pmatrix}$$

Using the Taylor expansion of $\frac{1}{1+x} = 1 - x + O(x^2)$, we can now form the inverse of $\mathbf{O}^{\frac{1}{2}} \approx \mathbf{1} + \Delta\mathbf{A}$. Thus

$$\mathbf{O}^{-\frac{1}{2}} = \begin{pmatrix} 1 & -\frac{1}{2}\Delta \\ -\frac{1}{2}\Delta & 1 \end{pmatrix} + O(\Delta^2)$$

Now we transform the Hamiltonian into the new basis

$$\begin{aligned}
\mathbf{H}' &= \mathbf{O}^{-\frac{1}{2}}\mathbf{H}\mathbf{O}^{-\frac{1}{2}} \\
&= \begin{pmatrix} 1 & -\frac{1}{2}\Delta \\ -\frac{1}{2}\Delta & 1 \end{pmatrix} \begin{pmatrix} \bar\epsilon_1 & t \\ t & \bar\epsilon_2 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2}\Delta \\ -\frac{1}{2}\Delta & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & -\frac{1}{2}\Delta \\ -\frac{1}{2}\Delta & 1 \end{pmatrix} \begin{pmatrix} \bar\epsilon_1 - \frac{t}{2}\Delta & t - \frac{\bar\epsilon_1}{2}\Delta \\ t - \frac{\bar\epsilon_2}{2}\Delta & \bar\epsilon_2 - \frac{t}{2}\Delta \end{pmatrix} \\
&= \begin{pmatrix} \bar\epsilon_1 - t\Delta & t - \frac{\bar\epsilon_1+\bar\epsilon_2}{2}\Delta \\ t - \frac{\bar\epsilon_1+\bar\epsilon_2}{2}\Delta & \bar\epsilon_2 - t\Delta \end{pmatrix} + O(\Delta^2)
\end{aligned}$$

Now we have, to first order in Δ an orthonormal basisset and the corresponding Hamiltonian.

B.2 Approximate Hamiltonian with the Wolfsberg formula

Now we use the Wolfsberg formula, which claims that the hopping matrix elements are proportional to the overlap matrix elements Δ. Thus, to first order in Δ, the product $t\Delta$ vanishes. We obtain:

$$\begin{aligned}
\mathbf{H}' &= \begin{pmatrix} \bar\epsilon_1 & k\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\Delta - \frac{\epsilon_1+\epsilon_2}{2}\Delta \\ k\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\Delta - \frac{\epsilon_1+\epsilon_2}{2}\Delta & \epsilon_2 \end{pmatrix} + O(\Delta^2) \\
&= \begin{pmatrix} \bar\epsilon_1 & (k-1)\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\Delta \\ (k-1)\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\Delta & \bar\epsilon_2 \end{pmatrix} + O(\Delta^2)
\end{aligned}$$

B.3 Origin of the Wolfsberg formula

Consider two atomic orbitals $|\chi_1\rangle$ and $|\chi_2\rangle$, centered on different atoms, which form a bond. The potential in the molecule shall be $\hat v$. The two orbitals obey the Schrödinger equation

$$\left(\frac{\hat p^2}{2m} + \hat v_i - \bar\epsilon_i\right)|\chi_i\rangle = 0$$

Let us try to work out an approximate expression for the hopping parameter t.

$$\begin{aligned}
t &= \langle\chi_1|\frac{\hat p^2}{2m} + \hat v|\chi_2\rangle \\
&= \frac{1}{2}\langle\chi_1|\frac{\hat p^2}{2m} + \hat v_1|\chi_2\rangle + \frac{1}{2}\langle\chi_1|\frac{\hat p^2}{2m} + \hat v_2|\chi_2\rangle + \langle\chi_1|\hat v - \frac{1}{2}\left(\hat v_1 + \hat v_2\right)|\chi_2\rangle \\
&= \frac{\bar\epsilon_1 + \bar\epsilon_2}{2}\langle\chi_1|\chi_2\rangle + \langle\chi_1|\hat v - \frac{1}{2}\left(\hat v_1 + \hat v_2\right)|\chi_2\rangle
\end{aligned}$$

This is still an exact expression. Now we make some approximations: Because the product of the two orbitals is only appreciable in the bonding region, we can replace the potentials by their values in the bond center $\vec r_b$.

$$t \approx \frac{\bar\epsilon_1 + \bar\epsilon_2}{2}\langle\chi_1|\chi_2\rangle + \left(v(\vec r_b) - \frac{1}{2}\left(v_1(\vec r_b) + v_2(\vec r_b)\right)\right)\underbrace{\Delta}_{\langle\chi_1|\chi_2\rangle}$$

This already suggests that the hopping matrix element t is approximately proportional to the overlap.

The second approximation is to assume that the potentials are approximately proportional to the energies, that is

$$\underbrace{v(\vec{r}_b) - \frac{1}{2}\left(v_1(\vec{r}_b) + v_2(\vec{r}_b)\right)}_{<0} \sim \frac{\bar{\epsilon}_1 + \bar{\epsilon}_2}{2}$$

I have not found a convincing justification for this latter approximation, but that is what it is.

Appendix C

Two-center bond with overlap

Let us first consider the hydrogen atom. We assume that the Hamilton matrix is given by

$$\mathbf{H} = \begin{pmatrix} \epsilon_1^0 & t \\ t & \epsilon_2^0 \end{pmatrix}$$

and the overlap matrix is

$$\mathbf{O} = \begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix}$$

In order to obtain the eigenvalues, we need to determine the zeros of the determinant

$$\begin{aligned} \det|\mathbf{H} - \epsilon\mathbf{O}| &= 0 \\ \Rightarrow \det\begin{vmatrix} \epsilon_1^0 - \epsilon & t - \epsilon s \\ t - \epsilon s & \epsilon_2^0 - \epsilon \end{vmatrix} &= 0 \\ \Rightarrow (\epsilon_1^0 - \epsilon)(\epsilon_2^0 - \epsilon) - (t - \epsilon s)^2 &= 0 \end{aligned}$$

C.1 Degenerate interaction

Let us first consider the non-degenerate case, that is

$$\epsilon_1^0 = \epsilon_2^0 =: \epsilon^0$$

The secular equation has now the form

$$\begin{aligned} (\epsilon^0 - \epsilon)^2 - (t - \epsilon s)^2 &= 0 \\ \epsilon - \epsilon^0 &= \pm(t - \epsilon s) \\ \epsilon - \epsilon^0 &= \pm(t - (\epsilon - \epsilon^0)s - \epsilon^0 s) \\ (1 \pm s)(\epsilon - \epsilon^0) &= \pm(t - \epsilon^0 s) \\ \epsilon &= \epsilon^0 \pm \frac{t - \epsilon^0 s}{1 \pm s} \end{aligned}$$

Thus we obtain the two energies as

$$\begin{aligned} \epsilon_- &= \epsilon^0 - \frac{\epsilon^0 s - t}{1 + s} \\ \epsilon_+ &= \epsilon^0 + \frac{\epsilon^0 s - t}{1 - s} \end{aligned}$$

We see that the energies split into a bonding orbital at lower energy ϵ_0 and an anti-bonding orbital at higher energy ϵ_+. However the lowering of the bonding orbital is smaller than the raise of the anti-bonding orbital.

In order to understand the consequences let us employ the Wolfsberg-Helmholtz formula[4]

$$t = (\epsilon_1^0 + \epsilon_2^0)s$$

The atomic orbital is usually negative, that is $\epsilon_0 < 0$. By choosing sensible sign conventions for the orbital we can make the overlap between the neighboring orbitals positive, that is $s > 0$. For reasonable systems $\epsilon_0 s - t$ is positive.

C.2 Nondegenerate interaction

$$\begin{aligned}
(\epsilon_1^0 - \epsilon)(\epsilon_2^0 - \epsilon) - (t - \epsilon s)^2 &= 0 \\
(1 - s^2)\epsilon^2 - (\epsilon_1^0 + \epsilon_2^0 - 2ts)\epsilon + \epsilon_1^0\epsilon_2^0 - t^2 &= 0 \\
\epsilon^2 - \frac{\epsilon_1^0 + \epsilon_2^0 - 2ts}{1 - s^2}\epsilon + \frac{\epsilon_1^0\epsilon_2^0 - t^2}{1 - s^2} &= 0 \\
\epsilon = \frac{\epsilon_1^0 + \epsilon_2^0 - 2ts}{2(1 - s^2)} \pm \sqrt{\left(\frac{\epsilon_1^0 + \epsilon_2^0 - 2ts}{2(1 - s^2)}\right)^2 - \frac{\epsilon_1^0\epsilon_2^0 - t^2}{1 - s^2}} \\
= \frac{1}{1 - s^2}\left[\frac{\epsilon_1^0 + \epsilon_2^0}{2} - ts \pm \sqrt{\left(\frac{\epsilon_1^0 + \epsilon_2^0}{2} - ts\right)^2 - \left(\epsilon_1^0\epsilon_2^0 - t^2\right)\left(1 - s^2\right)}\right] \\
= \frac{1}{1 - s^2}\left[\frac{\epsilon_1^0 + \epsilon_2^0}{2} - ts \pm \sqrt{\left(\frac{\epsilon_1^0 + \epsilon_2^0}{2}\right)^2 - 2ts\frac{\epsilon_1^0 + \epsilon_2^0}{2} + t^2s^2 - \epsilon_1^0\epsilon_2^0 + t^2 + s^2\epsilon_1^0\epsilon_2^0 - s^2t^2}\right] \\
= \frac{1}{1 - s^2}\left[\frac{\epsilon_1^0 + \epsilon_2^0}{2} - ts \pm \sqrt{\left(\frac{\epsilon_1^0 - \epsilon_2^0}{2}\right)^2 - 2ts\frac{\epsilon_1^0 + \epsilon_2^0}{2} + t^2 + s^2\epsilon_1^0\epsilon_2^0}\right]
\end{aligned}$$

C.3 Degenerate interaction

Let us first consider the non-degenerate case, that is

$$\epsilon_1^0 = \epsilon_2^0 =: \epsilon^0$$

We obtain

$$\begin{aligned}
\epsilon &= \frac{1}{1 - s^2}\left[\epsilon^0 - ts \pm \sqrt{t^2 - 2ts\epsilon^0 + s^2(\epsilon^0)^2}\right] \\
&= \frac{1}{1 - s^2}\left[\epsilon^0 - ts \pm \sqrt{(t - s\epsilon^0)^2}\right] \\
&= \frac{1}{1 - s^2}\left[\epsilon^0 - ts \pm \left|t - s\epsilon^0)\right|\right]
\end{aligned}$$

Appendix D

Covalent bonds to second order

This section shows a simple method to estimate the net covalent bonding energy.

Consider a Hamiltonian $\hat{H}^{(0)}$ and a perturbation $\hat{H}^{(1)}$. We have in mind that the unperturbed Hamiltonian $\hat{H}^{(0)}$ is diagonal in a basisset of tight-binding orbitals and that $\hat{H}^{(1)}$ contains the hopping matrix elements, between tight-binding orbitals. The unperturbed Hamiltonian has orthonormal eigenstates $|\psi_n^{(0)}\rangle$ with energies $\epsilon_n^{(0)}$.

The first-order change of the wave function due to a perturbation $\hat{H}^{(1)}$ is obtained from the first-order terms of the Schrödinger equation as follows:

$$
\begin{aligned}
& \Big(\hat{H}^{(0)} - \epsilon_n^{(0)}\Big)|\psi_n^{(1)}\rangle + \Big(\hat{H}^{(1)} - \epsilon_n^{(1)}\Big)|\psi_n^{(0)}\rangle = 0 \\
\Rightarrow \quad |\psi_n^{(1)}\rangle &= \Big(\epsilon^{(0)} - \hat{H}^{(0)}\Big)^{-1}\Big(\hat{H}^{(1)} - \epsilon_n^{(1)}\Big)|\psi_n^{(0)}\rangle \\
&= \sum_m |\psi_m^{(0)}\rangle\langle\psi_m^{(0)}|\Big(\epsilon_n^{(0)} - \hat{H}^{(0)}\Big)^{-1}\Big(\hat{H}^{(1)} - \epsilon_n^{(1)}\Big)|\psi_n^{(0)}\rangle \\
&= \sum_m |\psi_m^{(0)}\rangle \frac{H_{m,n}^{(1)} - \epsilon_n^{(1)}\langle\psi_m^{(0)}|\psi_n^{(0)}\rangle}{\epsilon_n^{(0)} - \epsilon_m^{(0)}} \\
&= \sum_{m\neq n} |\psi_m^{(0)}\rangle \frac{H_{m,n}^{(1)}}{\epsilon_n^{(0)} - \epsilon_m^{(0)}}
\end{aligned}
\tag{D.1}
$$

In the last step we imposed that the norm of the wave function remains conserved, so that $\langle\psi_m^{(0)}|\psi_n^{(1)}\rangle = 0$.

Furthermore, we will exploit that the states shall be orthonormal to any order, which implies that

$$
\langle\psi_m^{(0)} + \psi_m^{(1)} + \psi_m^{(2)} + \ldots|\psi_n^{(0)} + \psi_n^{(1)} + \psi_n^{(2)} + \ldots\rangle = \delta_{m,n}
\tag{D.2}
$$

in every order or the perturbation. This implies in the first order

$$
\langle\psi_n^{(1)}|\psi_m^{(0)}\rangle + \langle\psi_n^{(0)}|\psi_m^{(1)}\rangle = 0 \quad \Rightarrow \quad \mathrm{Re}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle = 0
\tag{D.3}
$$

and in the second order

$$
\langle\psi_n^{(2)}|\psi_m^{(0)}\rangle + \langle\psi_n^{(1)}|\psi_m^{(1)}\rangle + \langle\psi_n^{(0)}|\psi_m^{(2)}\rangle = 0 \Rightarrow \langle\psi_n^{(2)}|\psi_n^{(0)}\rangle + \langle\psi_n^{(0)}|\psi_n^{(2)}\rangle = -\langle\psi_n^{(1)}|\psi_n^{(1)}\rangle
\tag{D.4}
$$

With the assumption that the occupations do not change, the energy to second order is obtained as

$$
\begin{aligned}
E &= \sum_n f_n \langle \psi_n^{(0)} + \psi_n^{(1)} + \psi_n^{(2)} + \ldots | \hat{H}^{(0)} + \hat{H}^{(1)} | \psi_n^{(0)} + \psi_n^{(1)} + \psi_n^{(2)} + \ldots \rangle \\
&= \sum_n f_n \underbrace{\langle \psi_n^{(0)} | \hat{H}^{(0)} | \psi_n^{(0)} \rangle}_{\epsilon_n^{(0)}} \\
&+ \sum_n f_n \Big(\underbrace{\langle \psi_n^{(1)} | \hat{H}^{(0)} | \psi_n^{(0)} \rangle + \langle \psi_n^{(0)} | \hat{H}^{(0)} | \psi_n^{(1)} \rangle}_{\langle \psi_n^{(1)} | \psi_n^{(0)} \rangle \epsilon_n^{(0)} + \epsilon_n^{(0)} \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle = 0 \quad \text{(see Eq. D.3)}} + \langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle \Big) \\
&+ \sum_n f_n \Big(\underbrace{\langle \psi_n^{(2)} | \hat{H}^{(0)} | \psi_n^{(0)} \rangle + \langle \psi_n^{(0)} | \hat{H}^{(0)} | \psi_n^{(2)} \rangle}_{-\epsilon_n^{(0)} \langle \psi_n^{(1)} | \psi_n^{(1)} \rangle \quad \text{(see Eq. D.4)}} + \langle \psi_n^{(1)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle + \langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_n^{(1)} \rangle + \langle \psi_n^{(1)} | \hat{H}^{(0)} | \psi_n^{(1)} \rangle \Big) \ldots \\
&= \sum_n f_n \Big(\epsilon_n^{(0)} + \langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle \Big) + \sum_n f_n \Big(\langle \psi_n^{(1)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle + \langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_n^{(1)} \rangle + \langle \psi_n^{(1)} | \hat{H}^{(0)} - \epsilon_n^{(0)} | \psi_n^{(1)} \rangle \Big) + \ldots \\
&= \sum_n f_n \Big(\epsilon_n^{(0)} + \langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle \Big) + \sum_n f_n \sum_{m; m \neq n} \Big(\frac{H_{n,m}^{(1)}}{\epsilon_n^{(0)} - \epsilon_m^{(0)}} \langle \psi_m^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle + \langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_m^{(0)} \rangle \frac{H_{m,n}^{(1)}}{\epsilon_n^{(0)} - \epsilon_m^{(0)}} \\
&\quad + \frac{H_{n,m}^{(1)}}{\epsilon_n^{(0)} - \epsilon_m^{(0)}} \underbrace{\langle \psi_m^{(0)} | \hat{H}^{(0)} - \epsilon_n^{(0)} | \psi_m^{(0)} \rangle}_{\epsilon_m^{(0)} - \epsilon_n^{(0)}} \frac{H_{m,n}^{(1)}}{\epsilon_n^{(0)} - \epsilon_m^{(0)}} \Big) + \ldots \\
&= \sum_n f_n \Big(\epsilon_n + \langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle \Big) + \sum_n f_n \Big(\sum_{m; m \neq n} \frac{\Big| H_{n,m}^{(1)} \Big|^2}{\epsilon_n^{(0)} - \epsilon_m^{(0)}} \Big) + \ldots
\end{aligned}
\tag{D.5}
$$

Thus we obtain to second order the energy change

SECOND ORDER BONDING

For an already insulating system, for which the occupations remain unchanged, the energy gain due to the hopping matrix elements to second order are obtained as

$$
\Delta E = \sum_n f_n \Big(\epsilon_n + \Delta \epsilon_n \Big) \tag{D.6}
$$

with the level shifts

$$
\Delta \epsilon_n = H_{n,n}^{(1)} + \sum_{m; m \neq n} \frac{\Big| H_{m,n}^{(1)} \Big|^2}{\epsilon_m^{(0)} - \epsilon_n^{(0)}} \tag{D.7}
$$

The matrix elements are defined as

$$
H_{m,n}^{(1)} = \langle \psi_m^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle \tag{D.8}
$$

Thus the shift of the energy levels are obtained by summing up all level repulsions.
The second-order expression is valid when the perturbation between occupied and unoccupied levels is small compared to the already existing band gap. It breaks down when $|H_{n,m}^{(1)}| > |\epsilon_m - \epsilon_n|$. If the unperturbed energy levels are separated by a band gap, however, the total energy is only sensitive to level repulsions between filled and empty states. The requirement is then reduced to the limitation that the hopping matrix elements between filled and empty states must be smaller than the band gap.

This derivation suggests to first optimize the system using all onsite interactions, and to redistribute the electrons to form ions. If this system produces a band gap separating the filled and empty states, and if the hopping matrix elements accross the band gap are smaller than the band gap, the remaining bonding contribution can be obtained by perturbation theory. In this case the diagonal elements of the Hamiltonian form the zero'th order term $\hat{H}_0$, while the off-diagonal elements of the Hamiltonian, the hopping matrix elements, are treated as perturbation.

The Hamiltonian would have the form

$$\hat{H} = \sum_{\alpha} |\chi_\alpha\rangle \bar{\epsilon}_\alpha \langle\chi_\alpha| - \sum_{\alpha,\beta;\alpha\neq\beta} |\chi_\alpha\rangle t_{\alpha,\beta} \langle\chi_\beta| \tag{D.9}$$

where $|\chi_\alpha\rangle$ are the tight-binding orbitals, $\bar{\epsilon}_\alpha$ their energies, and $t_{\alpha,\beta}$ is the hopping matrix element between orbitals $|\chi_\alpha\rangle$ and $|\chi_\beta\rangle$.

Because, in this case, the perturbation vanishes on the diagonal, the first order term vanishes, and the change of the total energy is entirely fiven by the hopping matrix elements, which enter to second order.

APPROXIMATE COVALENT ENERGY AT FIXED OCCUPATIONS

In this special second order approximation, we obtain the total energy due to level shifts, i.e. for fixed occupations, in the form

$$\Delta E^{cov} \approx - \sum_{m,n;f_\alpha > f_\beta} |t_{\alpha,\beta}|^2 \frac{f_n - f_m}{\epsilon_m^{(0)} - \epsilon_n^{(0)}} \ . \tag{D.10}$$

The Hamiltonian is of the form Eq. D.9.

In this manner we can add up the approximate binding energy in a bond-by-bond manner.

The following is not worked out! This procedure works for ionic systems. One way to access covalent systems, where the preconditions of the above derivation are violated, is to start from a fictitious system with a negative on-site Coulomb interaction. This stabilized ionic systems. Then the hopping energy can be evaluated using the formula above. Finally one needs to correct again for the negative Coulomb interaction. This requires however the onsite occupations to be evaluated to second order.

Appendix E

Atomic structures

In this chapter I will keep the idealized atomic structure mentioned in the text. They are in the style of the input data of the CP-PAW code, so that they can copied from the pdf and easily be adapted. When doing so please ensure that you remove all non-printing characters.

E.1 fcc

```
!lattice T= 0.0 0.5 0.5 0.5 0.0 0.5 0.5 0.5 0.0 !end
!atom name='Ag1'  R=0.000 0.000 0.000  !end
```

E.2 bcc

```
!lattice T= -0.5 0.5 0.5 0.5 -0.5 0.5 0.5 0.5 -0.5 !end
!atom name='Fe1'  R=0.000 0.000 0.000  !end
```

E.3 Diamond structure

```
!lattice T= 0.0 0.5 0.5 0.5 0.0 0.5 0.5 0.5 0.0 !end
!atom name='C_1'  R=0.00 0.00 0.00  !end
!atom name='C_2'  R=0.25 0.25 0.25  !end
```

E.4 Rocksalt NaCl structure

```
!lattice T= 0.0 0.5 0.5 0.5 0.0 0.5 0.5 0.5 0.0 !end
!atom name='Na1'  R=0.00 0.00 0.00  !end
!atom name='Cl2'  R=0.50 0.50 0.50  !end
```

E.5 Fluorite structure

```
!lattice T= 0.0 0.5 0.5 0.5 0.0 0.5 0.5 0.5 0.0 !end
!atom name='Ca1'  R=0.00 0.00  0.00  !end
!atom name='F_1'  R=0.25 0.25  0.25  !end
!atom name='F_2'  R=0.75 0.75  0.75  !end
```

E.6 CsCl structure

```
  !lattice T= 1.0 0.0 0.0 0.0 1.0 0.0 0.0 0.0 1.0 !end
  !atom name=’Cs’  R=0.0 0.0 0.0 !end
  !atom name=’Cl’  R=0.5 0.5 0.5 !end
```

E.7 Spinel

56 atom non-primitive cubic unit cell of Spinel

```
  !lattice T= 2. 0. 0. 0. 2. 0. 0. 0. 2. !end
# plane 1
  !atom name=’O_1’  R=0.000 0.000 0.000  !end
  !atom name=’O_2’  R=1.000 0.000 0.000  !end
  !atom name=’O_3’  R=0.500 0.500 0.000  !end
  !atom name=’O_4’  R=1.500 0.500 0.000  !end
  !atom name=’O_5’  R=0.000 1.000 0.000  !end
  !atom name=’O_6’  R=1.000 1.000 0.000  !end
  !atom name=’O_7’  R=0.500 1.500 0.000  !end
  !atom name=’O_8’  R=1.500 1.500 0.000  !end
# plane 2
  !atom name=’O_9’  R=0.500 0.000 0.500  !end
  !atom name=’O_10’  R=1.500 0.000 0.500  !end
  !atom name=’O_11’  R=0.000 0.500 0.500  !end
  !atom name=’O_12’  R=1.000 0.500 0.500  !end
  !atom name=’O_13’  R=0.500 1.000 0.500  !end
  !atom name=’O_14’  R=1.500 1.000 0.500  !end
  !atom name=’O_15’  R=0.000 1.500 0.500  !end
  !atom name=’O_16’  R=1.000 1.500 0.500  !end
# plane 3
  !atom name=’O_17’  R=0.000 0.000 1.000  !end
  !atom name=’O_18’  R=1.000 0.000 1.000  !end
  !atom name=’O_19’  R=0.500 0.500 1.000  !end
  !atom name=’O_20’  R=1.500 0.500 1.000  !end
  !atom name=’O_21’  R=0.000 1.000 1.000  !end
  !atom name=’O_22’  R=1.000 1.000 1.000  !end
  !atom name=’O_23’  R=0.500 1.500 1.000  !end
  !atom name=’O_24’  R=1.500 1.500 1.000  !end
# plane 4
  !atom name=’O_25’  R=0.500 0.000 1.500  !end
  !atom name=’O_26’  R=1.500 0.000 1.500  !end
  !atom name=’O_27’  R=0.000 0.500 1.500  !end
  !atom name=’O_28’  R=1.000 0.500 1.500  !end
  !atom name=’O_29’  R=0.500 1.000 1.500  !end
  !atom name=’O_30’  R=1.500 1.000 1.500  !end
  !atom name=’O_31’  R=0.000 1.500 1.500  !end
  !atom name=’O_32’  R=1.000 1.500 1.500  !end
# plane1
  !atom name=’Al_1’  R=0.500 0.000 0.000  !end
  !atom name=’Al_2’  R=1.000 0.500 0.000  !end
  !atom name=’Al_3’  R=1.500 1.000 0.000  !end
  !atom name=’Al_4’  R=0.000 1.500 0.000  !end
# plane2
```

```
  !atom name='Al_5'  R=1.000 0.000 0.500  !end
  !atom name='Al_6'  R=0.500 0.500 0.500  !end
  !atom name='Al_7'  R=0.000 1.000 0.500  !end
  !atom name='Al_8'  R=1.500 1.500 0.500  !end
# plane3
  !atom name='Al_1'  R=1.500 0.000 1.000  !end
  !atom name='Al_2'  R=0.000 0.500 1.000  !end
  !atom name='Al_3'  R=0.500 1.000 1.000  !end
  !atom name='Al_4'  R=1.000 1.500 1.000  !end
# plane4
  !atom name='Al_9'  R=0.000 0.000 1.500  !end
  !atom name='Al_10' R=1.500 0.500 1.500  !end
  !atom name='Al_11' R=1.000 1.000 1.500  !end
  !atom name='Al_12' R=0.500 1.500 1.500  !end
# Mg
  !atom name='Mg1'  R=1.750 0.2500 0.250  !end
  !atom name='Mg2'  R=0.750 1.2500 0.250  !end
  !atom name='Mg3'  R=1.250 0.7500 0.750  !end
  !atom name='Mg4'  R=0.250 1.7500 0.750  !end
  !atom name='Mg5'  R=0.750 0.2500 1.250  !end
  !atom name='Mg6'  R=1.750 1.2500 1.250  !end
  !atom name='Mg7'  R=0.250 0.7500 1.750  !end
  !atom name='Mg8'  R=1.250 1.7500 1.750  !end
```

Appendix F

Reading

- Walter A Harrison has written a series of books describing the properties of materials on the basis of the tight-binding theory. “Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond” ”Elementary Electronic Structure”
- a magnificent resource is the book “Orbital interactions in Chemistry” by Albright, Burdett and Whangbo. It provides a good feel for orbitals and how they relate to atomic structures. Mostly for organic and organometallic chemistry.
- An easy to read introduction to simple molecular orbital theory of solids and surfaces is the book “Solids and Surfaces: A chemist’s View of Bonding in extended Structures” by R. Hoffmann.[49] The book covers basically the material of the article[50].
- The book “Computational Chjemistry of Solid state Materials” by R. Dronskowski[51] provides a critical overview over simple empirical rules of Inorganic chemistry. In addition it contains a description of computational simulation techniques, which is complemented with own applications.
- Chemogenesis webbook: `http://www.meta-synthesis.com/webbook.html`

Appendix G

Dictionary

principal quantum number	Hauptquantenzahl

Appendix H

Greek Alphabet

A	α	alpha	N	ν	nu
B	β	beta	Ξ	ξ	ksi
Γ	γ	gamma	O	$o,$	omicron
Δ	δ	delta	Π	π, ϖ	pi
E	ϵ, ε	epsilon	P	ρ, ϱ	rho
Z	ζ	zeta	Σ	σ, ς	sigma
H	η	eta	T	τ	tau
Θ	θ, ϑ	theta	Υ	υ	upsilon
I	ι	iota	Φ	ϕ, φ	phi
K	κ	kappa	X	χ	chi
Λ	λ	lambda	Ψ	ϕ	phi
M	μ	mu	Ω	ω	omega

Appendix I

Philosophy of the ΦSX Series

In the ΦSX series, I tried to implement what I learned from the feedback given by the students which attended the courses and that relied on these books as background material.

The course should be **self-contained**. There should not be any statements "as shown easily..." if, this is not true. The reader should not need to rely on the author, but he should be able to convince himself, if what is said is true. I am trying to be as complete as possible in covering all material that is required. The basis is the mathematical knowledge. With few exceptions, the material is also developed in a sequence so that the material covered can be understood entirely from the knowledge covered earlier.

The derivations shall be **explicit**. The novice should be able to step through every single step of the derivation with reasonable effort. An advanced reader should be able to follow every step of the derivations even without paper and pencil.

All **units** are explicit. That is, formulas contain all fundamental variables, which can be inserted in any desirable unit system. Expressions are consistent with the SI system, even though I am quoting some final results in units, that are common in the field.

The equations that enter a specific step of a derivation are noted as **hyperlinks** ontop of the equation sign. The experience is that the novice does not immediately memorize all the material covered and that he is struggling with the math, so that he spends a lot of time finding the rationale behind a certain step. This time is saved by being explicit about it. The danger that the student gets dependent on these indications, is probably minor, as it requires some effort for the advanced reader to look up the assumptions, an effort he can save by memorizing the relevant material.

Important results and equations are highlighted by including them in **boxes**. This should facilitate the preparations for examinations.

Portraits of the key researchers and short biographical notes provide independent associations to the material. A student may not memorize a certain formula directly, but a portrait. From the portrait, he may associate the correct formula. The historical context provides furthermore an independent structure to organize the material.

The two first books are in german (That is the intended native language) in order to not add complications to the novice. After these first books, all material is in English. It is mandatory that the student masters this language. Most of the scientific literature is available only in English. English is currently the language of science, and science is absolutely dependent on international contacts.

I tried to include many graphs and figures. The student shall become used to use all his senses in particular the **visual sense**.

I have slightly modified the selection of the material commonly tought in most courses. Some topics, which I consider of mostly historical relevance I have removed. Others such as the Noether theorem, I have added. Some, like chaos, stochastic processes, etc. I have not added yet.

Appendix J

About the Author

Prof. Dr. rer. nat Peter E. Blöchl studied physics at Karlsruhe University of Technology in Germany. Subsequently he joined the Max Planck Institutes for Materials Research and for Solid state Research in Stuttgart, where he developed of electronic-structure methods related to the LMTO method and performed first-principles investigations of interfaces. He received his doctoral degree in 1989 from the University of Stuttgart.

Following his graduation, he joined the renowned T.J. Watson Research Center in Yorktown Heights, NY in the US on a World-Trade Fellowship. In 1990 he accepted an offer from the IBM Zurich Research Laboratory in Ruschlikon, Switzerland, which had just received two Nobel prices in Physics (For the Scanning Tunneling Microscope in 1986 and for the High-Temperature Superconductivity in 1987). He spent the summer term 1995 as visiting professor at the Vienna University of Technology in Austria, from where was later awarded the habilitation in 1997. In 2000, he left the IBM Research Laboratory after a 10-year period and accepted an offer to be professor for theoretical physics at Clausthal University of Technology in Germany. Since 2003, Prof. Blöchl is member of the Braunschweigische Wissenschaftliche Gesellschaft (Academy of Sciences).

The main thrust of Prof. Blöchl's research is related to ab-initio simulations, that is, parameter-free simulation of materials processes and molecular reactions based on quantum mechanics. He developed the Projector Augmented Wave (PAW) method, one of the most widely used electronic structure methods to date. This work has been cited over 35,000 times.[1] It is among the 100 most cited scientific papers of all times and disciplines[2] , and it is among the 10 most-cited papers out of more than 500,000 published in the 120-year history of Physical Review. [3] Next to the research related to simulation methodology, his research covers a wide area from biochemistry, solid state chemistry to solid state physics and materials science. Prof. Blöchl contributed to 8 Patents and published about 100 research publications, among others in well-known Journals such as "Nature". The work of Prof. Blöchl has been cited over 48,000 times, and he has an H-index of 44.[4]

---

[1] Researcher ID: B-3448-2012

[2] R. van Noorden, B. Maher and R. Nuzzo, Nature 514, 550 (2014)

[3] Oct. 15, 2014, search in the Physical Review Online Archive with criteria "a-z".

[4] Researcher ID: B-3448-2012

Bibliography

[1] T.A. Albright, J.K. Burdett, and M.-H. Whangbo. Orbital interactions in Chemistry. John Wiley and Sons, New York, 1985.

[2] F. Bloch. Über die Quantenmechanik der Elektronen in Kristallgittern. Zeitschr. Phys., 52:555, 1929.

[3] J.C. Slater and G.F. Koster. Simplified LCAO method for the periodic potential problem. Phys. Rev., 94:1498, 1954.

[4] Max Wolfsberg and Lindsay Helmholz. The spectra and electronic structure of the tetrahedral ions MnO_4^-, CrO_4^{2-}, and ClO_4^-. J. Chem. Phys, 20:837, 1952.

[5] K. Fukui. Nobel Lectures Chemistry, 1981-1990, chapter The role of frontier orbitals in chemical reactions. World Scientific, Singapore, 1982.

[6] P.E. Blöchl and J.H. Stathis. Hydrogen electrochemistry and stress-induced leakage current in silica. Phys. Rev. Lett, 83:372, 1999.

[7] P.E. Blöchl. First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen. Phys. Rev. B, 62:6158, 2000.

[8] W. Harrison. The physics of solid state chemistry. Advances in Solid State Physics, 17:135, 1977.

[9] S. Froyen and W.A. Harrison. Elementary prediction of linear combination of atomic orbital matrix elements. Phys. Rev. B, 20:2420, 1979.

[10] W.A. Harrison. Theory of the two-center bond. Phys. Rev. B, 27:3592, 1983.

[11] M. van Schilfgaarde and W.A. Harrison. Theory of the multicenter bond. Phys. Rev. B, 33: 2653, 1986.

[12] Konny Schmitt, Claudia Stückl, Herbert Ripplinger, and Barbara Albert. Crystal and electronic structure of BaB_6 in comparison with CaB_6 and molecular $[B_6H_6]^{2-}$. Solid State Sciences, 3 (3):321 – 327, 2001. ISSN 1293-2558. doi: https://doi.org/10.1016/S1293-2558(00)01091-8. URL `http://www.sciencedirect.com/science/article/pii/S1293255800010918`.

[13] R. Dronskowski and P.E. Blöchl. Crystal orbital Hamilton populations (COHP). Energy-resolved visualization of chemical bonding in solids based on density-functional calculations. J. Phys. Chem., 97:8617, 1993.

[14] P. E. Blöchl. ΦSX: Introduction to Solid State Theory. Blöchl, 2017. URL `http://www2.pt.tu-clausthal.de/atp/phisx.html`.

[15] R.S. Mulliken. A new electroaffinity scale; together with data on valence states and on valence ionization potentials and electron affinities. J. Chem. Phys., 2:782, 1934.

[16] H. A. Jahn and E. Teller. Stability of polyatomic molecules in degenerate electronic states. i. orbital degeneracy. Proc. R. Soc. London A, 1937.

[17] Linus Pauling. The nature of the chemical bond. III. The transition from one extreme bond type to another. J. Am. Chem. Soc., 54:988, 1932.

[18] H. Räbiger, S. Lany, and A. Zunger. Charge self-regulation upon changing the oxidation state of transition metals in insulators. Nature, 453:763, 2008.

[19] R.G. Parr and R.G. Pearson. Absolute hardness: Companion parameter to absolute electronegativity. J. Am. Chem. Soc., 105:7512, 1983.

[20] J.F. Janak. Proof that $\partial e/\partial n_i = \epsilon_i$ in density functional theory. Phys. Rev. B, 18:7165, 1978.

[21] Linus Pauling. The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms. J. Am. Chem. Soc., 54:3570, 1932.

[22] R.G. Pearson. Hard and soft acids and bases. J. Am. Chem. Soc., 85:3533, 1963.

[23] P.K. Chattaraj, H. Lee, and R.G. Parr. HSAB principle. J. Am. Chem. Soc., 113:1855, 1991.

[24] P.W. Ayers, R.G. Parr, and R.G. Pearson. Elucidating the hard/soft acid/base principle: A perspective based on half-reactions. J. Chem. Phys., 124:194107, 2006.

[25] P.P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. Ann. Phys., 64: 253, 1921. doi: 10.1002/andp.19213690304.

[26] Leslie Glasser. Solid-state energetics and electrostatics: Madelung constants and madelung energies. Inorg. Chem., 51:2420, 2012.

[27] L. Pauling. The nature of the chemical bond and the structure of molecules and crystals. Cornell University Press, 1960.

[28] M. Born and J.E. Mayer. Zur gittertheorie der ionenkristalle. Z. Phys., 75:1, 1932.

[29] A.F. Kapustinskii. Lattice energy of ionic crystals. Quarterly Reviews, Chem. Soc., 10:283, 1956.

[30] W.B. Jensen. The origin of the ionic-radius ratio rules. J. Chem. Educ., 87:587, 2010.

[31] G.F. Huettig. Notiz zur geometrie der koordinationszahl. Z. anorg. Chem., 114:24, 1920.

[32] Linus Pauling. The principles determining the structure of complex ionic crystals. J. Am. Chem. Soc., 51:1010, 1929.

[33] J.I. Gersten and F.W. Smith. The Physics and Chemistry of Materials. John Wiley and Sons, Inc., 2001.

[34] S.H.Pine. Organic Chemistry. McGraw Hill, 1987.

[35] J. Huheey, E.A. Keiter, and R.L. Keiter. Inorganic Chemistry: Principles of Structure and Reactivity. Harper Collins, New York, 1993.

[36] Michael J. Mehl, David Hicks, Cormac Toher, Ohad Levy, Robert M. Hansone, Gus Hartf, and Stefano Curtarolo. The AFLOW library of crystallographic prototypes. Arxiv, 1607:02532, 2016.

[37] David Hicks, Michael J. Mehl, Eric Gossett, Cormac Toher, Ohad Levy, Robert M. Hansone, Gus Hart, and Stefano Curtarolo. The aflow library of crystallographic prototypes: Part 2. Arxiv, 1806:07864, 2018.

[38] Fredrik Lindberg. Studies of Oxygen Deficient Complex Cobaltates with Perovskite Related Structures. PhD thesis, Physical, Inorganic and Structural Chemistry, Stockholm University, 2006.

[39] E. O. Wollan and W. C. Koehler. Neutron diffraction study of the magnetic properties of the series of perovskite-type compounds $[(1-x)\text{La}, x\text{Ca}]\text{Mno}_3$. Phys. Rev., 100:545–563, Oct 1955. doi: 10.1103/PhysRev.100.545. URL `https://link.aps.org/doi/10.1103/PhysRev.100.545`.

[40] J.C. Slater. The Lorentz correction in barium titanate. Phys. Rev., 78:748, 1950.

[41] R.E. Cohen. Origin of ferroelectricity in perovskite oxides. Nature, 358:136, 1992.

[42] A.F. Wells. The structuresof crystals. Solid State Physics, 7:425, 1958.

[43] K.K. Murata and S. Doniach. Theory of magnetic fluctuations in itinerant ferromagnets. Phys. Rev. Lett., 29:285, 1972.

[44] Peter Mohn and E.P. Wohlfarth. The Curie temperature of the ferromagnetic transition metals and their compounds. J. Phys. F: Met. Phys., 17:2421, 1987.

[45] C. Zener. Interaction between the d shells in the transition metals. Phys. Rev., 81:440, 1951.

[46] H.A. Kramers. L'interaction entre les atomes magnetogenes dans un cristal paramagnetique. Physica, 1:191, 1934.

[47] P.W Anderson. New approach to the theory of superexchange interactions. Phys. Rev., 115:2, 1959.

[48] Clarence Zener. Interaction between the *d*-shells in the transition metals. ii. ferromagnetic compounds of manganese with perovskite structure. Phys. Rev., 82:403–405, May 1951. doi: 10.1103/PhysRev.82.403. URL `https://link.aps.org/doi/10.1103/PhysRev.82.403`.

[49] R. Hoffmann. Solids and Surfaces: A Chemist's View of Bonding in Extended Structures. Wiley-VCH, 1988.

[50] R. Hoffmann. A chemical and theoretical way to look at bonding on surfaces. Rev. Mod. Phys., 60:601, 1988.

[51] R. Dronskowski. Computational Chemistry of Solid State Materials: A Guide for Material Scientists, Chemists, Physicists and others. Wiley VCH, 2005.

Index

angular-momentum quantum number, 61
anti-bonding state, 20

band structure, 84
Bloch theorem, 84
block diagonal, 33
bond
  ionic, 21
bonding state, 20
Born repulsion, 120
Born-Mayer equation, 121
Boson, 14
bra, 13
bracket notation, 13
Bravais lattice, 133
Brillouin zone, 96

characteristic polynomial, 18
chemical potential, 114
COHP, 83
commutating operations, 52
complete neglect of overlap, 17
COOP, 83
core repulsion, 105
Coulomb integral, 172
Crystal Orbital Hamilton Population, 83
Crystal Orbital Overlap Population, 83

delta function, 13
density of states
  total, 83
density functional theory, 15
density of states
  projected, 83
Density of States, 26
Density of states
  per volume, 86
Dirac's bracket notation, 13
Dirac's delta function, 13
dispersion relation
  linear chain, 26
double bond, 48
double exchange, 173

e_g, 73
eigenstate, 12
eigenvalue, 12
eigenvectors, 16
electron affinity, 106, 116
electronegativity, 107, 115
electronegativity
  absolute, 115
  Mulliken, 115
electronegativity
  Pauling, 115
electrophile, 23, 75
energy spectrum, 12
entartet, 73
ethene, 48
Ewald sum, 120
exchange coupling, 172
exchange splitting, 167
expectation value, 12
expectation value
  in bracket notation, 13
extended zone scheme, 95

Fermi-level, 22
Fermion, 14
free electron gas, 81
Frontier orbitals, 75

Hamilton function, 11
Hamilton operator, 12
  bracket notation, 13
Hamilton's equations, 11
hardness, 114, 116
Hartree-Fock approximation, 170
heat of dissociation, 122
HOMO, 23
hopping matrix element, 16
HSAB principle, 117
Hund's rule, 165
hybridization, 21

ionic bond, 21
ionization potential, 106, 116

Jahn-Teller theoem, 110
Janack's theorem, 114
jellium model, 81

Kapustinskii equation, 121
ket, 13

LAPACK, 16
Laplace-Runge-Lenz vector, 103
lattice constant, 89
lattice vector
    general, 89
    primitive, 89
    reciprocal, 89
LUMO, 23

Madelung constant, 120
magnetic quantum number, 61
martensitic transition, 150
methane, 52
momentum operator, 11
momentum operator
    in bracket notation, 13

node plane, 21
normalization condition, 12
nucleophile, 23, 75
number of states, 84

Observable, 11
oxidation state, 108, 113

Pauli principle, 14
periodic, 89
periodic boundary condition, 25
periodic zone scheme, 96
position operator, 13
positions
    relative, 89
promotion energy, 67

quantum numbers, 35
quasi crystal, 133

radical, 23
Reciprocal lattice vectors, 90
reduced zone scheme, 95

scalar product, 13
    in bracket notation, 13
Schrödinger equation
    time dependent, 12
    time independent, 12
Slater determinant, 14
Slater-Koster tables, 43
softness, 116
spherical harmonics, 61
spherical harmonics
    real, 62
spherical harmonics
    cubic, 62
spin, 14
star, 43
Stoner criterion, 167
symmetry, 34

t_{2g}, 73
three-center bond, 48
tight-binding method, 43
tight-binding model, 40
transformation operator, 34
triple, 73
triple bond, 48

unit cell
    primitive, 89

vacuum level, 116
Van Arkel-Ketelaar triangle, 117
van der Waal's distance, 130
van-der-Waal's bond, 129

Wigner Seitz cell, 96
Wolfsberg-Helmholtz formula, 17
work function, 116