

Serial: Quanta of orbitals

Today we'll look at how working with functions in quantum mechanics is in fact impractical, and how to simplify such a formulation. This part will be slightly more challenging to understand, so if you get a little bit lost, don't despair. The next episodes will again be easier to comprehend, and they won't be completely related to this one.

Oscillating in a quantum way

We're going to demonstrate everything on the linear harmonic oscillator. The system has a parabola-shaped potential

$$V(x) = \frac{1}{2}m\omega^2 x^2,$$

where ω is the frequency of the oscillator and m is the mass of the particle. If we substitute the relation for the frequency of the spring oscillation $\omega = \sqrt{k/m}$ we instead obtain the more familiar expression with stiffness k : $V = 1/2kx^2$. This system is one of the most important of all that we can solve analytically. (The other essential ones are the free particle and the hydrogen atom.) One of the reasons it is so important to us is the assumption that at any given potential, a particle will be in the vicinity of its minimum. The potential around the minimum can typically be approximated by a parabola. potential into the Schrödinger equation $H\psi = E\psi$ from the last episode, we get the equation

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2 x^2 \psi(x) = E\psi(x).$$

Although this could be solved directly, we will use it to present an entirely different idea. For simplicity, we can put $m = 1, \omega = 1$, and thus get rid of the need to rewrite them in the further equations.

Imagine someone reveals to you that the wave function of the ground state is a Gaussian function. We can verify that the function

$$\varphi_1(x) = e^{-x^2/2\hbar}$$

is a solution to the Schrödinger equation

$$\hat{H}\varphi_1(x) = \frac{\hbar}{2}e^{-x^2/2\hbar} = \frac{\hbar}{2}\varphi_1(x)$$

with energy $\hbar/2$. We can deal with normalization later, as it is not essential for energy calculation. And now, if we try to take the second function

$$\varphi_2(x) = x^2 e^{-x^2/2\hbar},$$

we find out that this function is not a solution to the Schrödinger equation

$$\hat{H}\varphi_2(x) = -\hbar^2 e^{-x^2/2\hbar} + 2\hbar x^2 e^{-x^2/2\hbar} = -\hbar^2 \varphi_1(x) + 5/2\hbar \varphi_2(x) \neq E\varphi_2(x).$$

Instead, applying the Hamiltonian to this function will add some contribution from the wave function of the ground state. However, there is nothing easier than trying to alter this function so that it satisfies the Schrödinger equation. This can be done simply by taking the sum of the 2 functions with unknown coefficients

$$\psi(x) = c_1\varphi_1(x) + c_2\varphi_2(x).$$

If we plug this form of the wave function into the Schrödinger equation, we get

$$\begin{aligned}\hat{H}\psi(x) &= c_1\hat{H}\varphi_1(x) + c_2\hat{H}\varphi_2(x) = \\ &= c_1\frac{\hbar}{2}\varphi_1(x) - c_2\hbar^2\varphi_1(x) + 5/2c_2\hbar\varphi_2(x) = \\ &= E(c_1\varphi_1(x) + c_2\varphi_2(x)).\end{aligned}\tag{1}$$

Since normalization is not necessary at this point, we can put $c_2 = 1$ and then just find the corresponding value of c_1 . This leaves us with an equation for the two unknowns, the coefficient c_1 and the energy E , from the original differential equation.

$$c_1\frac{\hbar}{2}\varphi_1(x) - \hbar^2\varphi_1(x) + 5/2\hbar\varphi_2(x) = E(c_1\varphi_1(x) + \varphi_2(x))$$

Looking at the coefficients of φ_2 , we see that the energy must equal $5/2\hbar$. Now we can easily calculate the coefficient c_1 .

$$c_1\frac{\hbar}{2}\varphi_1(x) - \hbar^2\varphi_1(x) = 5/2\hbar c_1\varphi_1(x),$$

We get $c_1 = \hbar/2$. This then gives us the wave function

$$\psi(x) = \hbar/2\varphi_1(x) + \varphi_2(x),$$

which, except for normalization, is exactly the result we can find in the textbooks. It's already the second excited state, because the first excited state is an odd function of x , while our two states are even. Since we have a symmetric potential, the Hamiltonian does not mix even and odd states. This is one of many cases where symmetry presents itself in quantum mechanics.

Let's try to give a second thought to what we've just done. Instead of solving a complicated differential equation for an unknown function, we took some (convenient!) set of fixed functions and just looked for the coefficients at their weighted sum (called a linear combination). By doing this, we converted the ugly differential equation into an algebraic equation, which is just more pleasant to solve. However, we must be careful to choose functions that describe the problem well. That way, we only need to take a relatively small amount of them to get a correct (or reasonably accurate approximation of the) solution. For example, if we wanted to describe the states of a linear harmonic oscillator using ordinary polynomials, you can surely guess that if we took any finite number of them, we would never get the wave function of the linear harmonic oscillator right. That's because the value of the exact wave function goes to zero for x going to infinity, and conversely, any polynomial goes to (positive or negative) infinity.

So far, however, we have determined the values of the coefficients a bit unsystematically, which we will now rectify. Let's go back a bit, to the equation (1). If we multiply the whole

thing by the function $\varphi_1(x)$ and make the integral with respect to x over the whole real axis, we get

$$\begin{aligned} & \int_{-\infty}^{\infty} (c_1\varphi_1(x)\hat{H}\varphi_1(x) + c_2\varphi_1(x)\hat{H}\varphi_2(x)) dx = \\ & = \int_{-\infty}^{\infty} E(c_1\varphi_1(x)^2 + c_2\varphi_1(x)\varphi_2(x)) dx, \end{aligned} \quad (2)$$

which we can easily modify to

$$\begin{aligned} & c_1 \int_{-\infty}^{\infty} (\varphi_1(x)\hat{H}\varphi_1(x)) dx + c_2 \int_{-\infty}^{\infty} (\varphi_1(x)\hat{H}\varphi_2(x)) dx = \\ & = E \left(c_1 \int_{-\infty}^{\infty} \varphi_1(x)^2 dx + c_2 \int_{-\infty}^{\infty} \varphi_1(x)\varphi_2(x) dx \right). \end{aligned} \quad (3)$$

Now we can notice that since $\varphi_1(x)$ and $\varphi_2(x)$ are fixed functions, we have all the necessary “ingredients” to compute the integrals and install their specific value. We’re not going to do that and instead, for generality, we’ll just denote them with letters

$$\begin{aligned} S_{11} &= \int_{-\infty}^{\infty} \varphi_1(x)^2 dx, \\ S_{12} &= \int_{-\infty}^{\infty} \varphi_1(x)\varphi_2(x) dx, \\ H_{11} &= \int_{-\infty}^{\infty} (\varphi_1(x)\hat{H}\varphi_1(x)) dx, \\ H_{12} &= \int_{-\infty}^{\infty} (\varphi_1(x)\hat{H}\varphi_2(x)) dx. \end{aligned}$$

The previous equation (3) then looks much nicer

$$H_{11}c_1 + H_{12}c_2 = E(S_{11}c_1 + S_{12}c_2).$$

Repeating this procedure, but multiplying the equation (2) by $\varphi_2(x)$ and reintegrating, we get an equation with the use of similar notation

$$H_{21}c_1 + H_{22}c_2 = E(S_{21}c_1 + S_{22}c_2).$$

Vectors and matrices

Now we need to take a little diversion to introduce a new, fancier notation. If you already have experience with matrices, feel free to skip this section. Unfortunately, this episode cannot be stretched, so the introduction will be very fast-paced, but if you want to learn more about matrices, there was, for example, an entire fifth episode of the 16th series devoted to them, which you can find on Fykos’ website.

I am sure you have already been introduced to vectors in school. If we have a coordinate system, we can write a vector in it as a column of numbers. Generalizing this concept for operations on vectors in space (such as rotation or extension), we get a table of numbers, which

we call a matrix. We will only use square matrices, but they can also be rectangular. The operation on vectors in these “tables of numbers” is carried over to the matrix multiplication operation. Its result is a vector, which at each position has the scalar product of the old vector with the corresponding row of the given matrix. Thus, the new vector has the scalar product of the first row with the original vector in the first position, and so on.

The application of the matrix \mathbb{A} on the vector \vec{x} which returns \vec{y} can be denoted as

$$\mathbb{A}\vec{x} = \vec{y}.$$

Similarly, we can multiply two matrices by each other – we break down the matrix in the product on the right into individual vectors, act on them with the left matrix, and then put these new vectors back together column by column into the resulting matrix. Multiplying two matrices is equivalent to composing two operations represented by those matrices. Beware, matrix multiplication is generally not commutative, $\mathbb{A}\mathbb{B} \neq \mathbb{B}\mathbb{A}$. Even stacking two representations is not commutative; you can see for yourself that if you rescale the image along the x-axis and then rotate it by 90° , you get something different than if you do it in reverse order. It’s going to be a little more illustrative as to why we’ve introduced this if we implement index notation. If we denote the i -th element of the vector \vec{x} as x_i and the matrix element \mathbb{A} in the i -th row and j -th column as A_{ij} , the action of the matrix on the vector goes to

$$\vec{y} = \mathbb{A}\vec{x} \rightarrow y_i = \sum_{j=1}^N A_{ij}x_j.$$

Well, that’s exactly what we have in our equation! Now we can declare the individual coefficients of c_i to be an (abstract!) vector, and thus we can rewrite the whole equation using matrices

$$\mathbb{H}\vec{c} = E\mathbb{S}\vec{c}.$$

The Hamiltonian matrix \mathbb{H} tells us how the Hamiltonian acts between the basis functions. The matrix \mathbb{S} has an even simpler interpretation. It basically tells us how much the individual basis functions overlap (we call it the overlapping matrix). Moreover, we can get rid of it easily, by choosing the basis functions such that the overlap of each function with itself is 1 and the overlap between any different basis functions is 0. In analogy with vectors in 3D space, we then say that we have orthogonal basis functions. The overlapping matrix then has a nice shape, with ones on the diagonal and zeros at all other positions. Such a matrix, when applied to any vector, will return the same vector, so we can drop it from the equation altogether. Then, for orthogonal basis functions, we have a much nicer form of the equation

$$\mathbb{H}\vec{c} = E\vec{c}.$$

This equation is exactly in the same form as the time-independent Schrödinger equation, but with matrices and vectors instead of functions and operators. The advantage is that for matrices, this type of equation is a standard problem, called diagonalization or finding eigenvalues and eigenvectors.

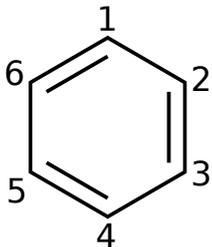
The eigenvector we are looking for changes only its magnitude by the action of the matrix, and not its orientation. Thus, after applying the matrix to the eigenvector, we get the same vector, just multiplied by some number. This is called the eigenvalue of the given vector. In principle, there are infinitely many eigenvectors, because by multiplying an eigenvector by any

non-zero number, we get an eigenvector with the same eigenvalue. So, we restrict ourselves to finding one of each “type”. Matrices usually have as many different eigenvectors as there are columns in the square matrix. Although for small matrices, such eigenvectors can be found analytically, it is beyond the limits of this text, and we are content to solve the problem numerically on the computer.

Counting molecules

We will now show how we could use a simple approximation to calculate the states of hydrocarbons with conjugated double bonds. This approximation is called the Hückel method, after Erich Hückel, who introduced it in 1930. As with the problem in the last part of the series, we restrict ourselves to p orbitals involved in double bonds, one for each carbon atom with a double bond. The basis functions are then just the individual p orbitals. Since the overlap between the p orbitals on different atoms is generally very small, we can neglect it and thus consider the entire basis set to be orthogonal. Then we only need to construct the Hamiltonian matrix. The latter is quite simple in this approximation. The value of the Hamiltonian for the interaction of an orbital with itself is just the energy of the electron in the p orbital of carbon with magnitude -11.4eV . This is denoted by α . Much more interesting are the elements of the Hamiltonian that are between orbitals on different atoms. If these atoms are not adjacent (i.e., not bonded), the interaction between them is neglected. The element of the Hamiltonian matrix will therefore be 0. If they are adjacent to each other, it has a non-zero value β , which is determined empirically, from experimental values for various properties of the molecules. The value of β varies depending on which molecules and experimental values are used to obtain it, but in general, its value is between -1eV and -3eV . Nonetheless, the Hückel method is used more for qualitative understanding, so the energies of the individual orbitals are usually expressed as multiples of β .

We’ll display an example of what the calculation would look like for a benzene molecule. We’re going to number the carbon atoms and make a table of what the elements of the Hamiltonian look like. On the diagonal of this table from the top left to the bottom right corner is *alpha*, because that is the Hamiltonian element for the orbital interacting with itself. The neighboring diagonals are β because the row and column have a value that differs by one, so (in our numbering) they are neighboring atoms. Finally, because our molecule is cyclic and the adjacent atoms are 1 and 6, we also get β in the upper right and lower left corners. Everywhere else, we’re left with zeros.



	1	2	3	4	5	6
1	α	β	0	0	0	β
2	β	α	β	0	0	0
3	0	β	α	β	0	0
4	0	0	β	α	β	0
5	0	0	0	β	α	β
6	β	0	0	0	β	α

This table is the very matrix of the Hamiltonian, and we can find its numbers directly.

Before that, it is sensible to rewrite this matrix to a form that is easier to calculate with

$$\mathbb{H} = \begin{pmatrix} \alpha & \beta & 0 & 0 & 0 & 0 & \beta \\ \beta & \alpha & \beta & 0 & 0 & 0 & \\ 0 & \beta & \alpha & \beta & 0 & 0 & \\ 0 & 0 & \beta & \alpha & \beta & 0 & \\ 0 & 0 & 0 & 0 & \beta & \alpha & \beta \\ \beta & 0 & 0 & 0 & \beta & \alpha & \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Note that the first matrix with prefactor α is unitary, so it acts on all vectors equally. Thus, its effect is just a shift of the energy of all states by α . Hence, we only need to diagonalize the second matrix (without β) and we get the individual energy levels in multiples of β . In this text, we will not attempt an analytical solution (even though it would be solvable thanks to the nice shape of this matrix) and will be satisfied with a numerical solution. We can use online tools to diagonalize the matrix, a well-known one is a www.wolframalpha.com, where we would use the command "eigenvalues0,1,0,0,0,1,1,0,1,0,0,0,0,1,0,1,0,0,0,0,1,0,1,0,"

{0,0,0,1,0,1},{1,0,0,0,1,0}}. We used a short Python script for the diagonalization.

```
#!/usr/bin/python
```

```
import numpy as np
```

```
N = 6
```

```
H = np.zeros((N, N)) #initialize the matrix of zeros NxN
```

```
for i in range(N-1): #run the whole matrix and set 1 next to the diagonal
```

```
    H[i, i+1]=1
```

```
    H[i+1, i]=1
```

```
#set 2 elements in the rows - index (N-1)
```

```
#because we are numbering from 0!
```

```
H[N-1,0]=1
```

```
H[0,N-1]=1
```

```
eig_values, eig_vector = np.linalg.eig(H) #custom diagonalization
```

```
print(eig_values) #print the result
```

After running this script, we get the eigenvalues $-2, -1, -1, 1, 1, 2$. The fact that there are some eigenvalues twice just means that we have two different eigenvectors with the same eigenvalue. These are also called degenerate eigenstates. But this is not the end of our calculations. Since we “forgot” the unit matrix and the prefactors α and β when we diagonalized the Hamiltonian, we now have to put them back in. That is not a problem, and we get eigenstates with energies $\alpha - 2\beta, \alpha - \beta, \alpha - \beta, \alpha + \beta, \alpha + \beta, \alpha + 2\beta$. Thus, we have 6 states (orbitals) onto which we must place 6 electrons in such a way as to minimize the energy. This means that we occupy the 3 states with the lowest energy twice. The resulting energy is then $6\alpha + 8\beta$. (Remember that both α and β are negative!)

However, the total energy is not something we’re actually interested in if only because in our approximation it doesn’t include the other electrons that don’t participate in double bonds

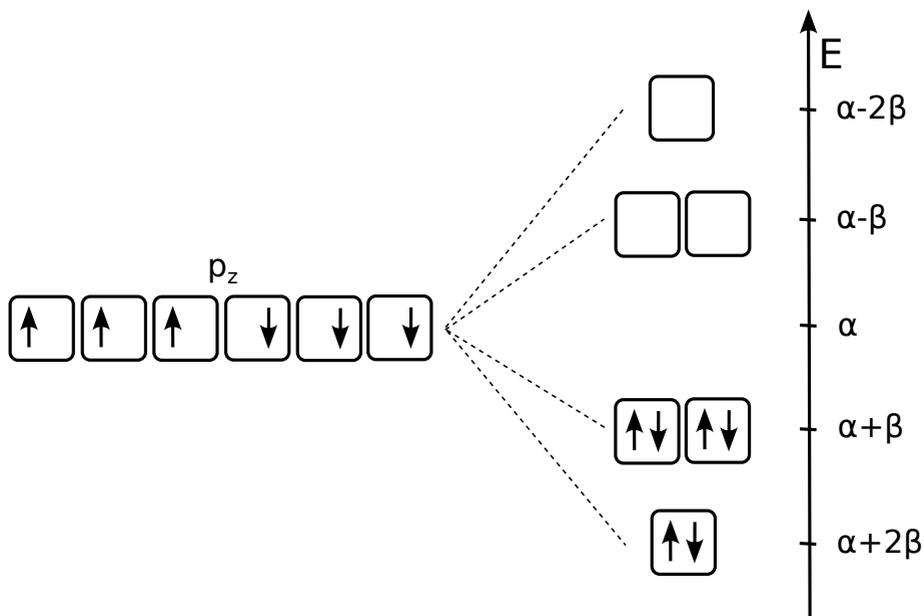


Fig. 1: The orbitals in the benzene molecule. On the left is a hypothetical state, where the p_z orbitals in benzene would not interact at all. The influence of interaction gives us the orbitals on the right, where the electrons naturally possess those with the lowest energy. The energy of the orbitals increases from bottom to top.

at all. So, let's try to extract some other properties that can be experimentally detected. The excitation energy, the energy required to excite from the highest occupied to the lowest virtual level, is -2β . After substituting a specific value $\beta = -3\text{ eV}$, we get excitation energy of 6 eV . Since both the highest occupied and the lowest unoccupied orbitals are twice degenerated, we are left with 4 different possible excitations. Although Hückel's model assigns them the same energy they have been experimentally confirmed to have 3 different values, namely 4.90 eV , 6.20 eV , and 6.94 eV (one of the energies corresponds to 2 excitations). We see that our rough approximation with a suitable empirical value of β provides quite good results.

If we calculate the ionization energy (the energy needed to pull an electron out of the highest occupied orbital), we get the negative energy of that orbital, $-\alpha - \beta$. After substitution, we will get more than 14 eV . You can see that we are starting to move away from the experimental value of 9.24 eV . If we try the same thing for the electron affinity, the energy released when an electron is added, we have $-\alpha + \beta$, a little over 8 eV , which is totally off the actual value of 1.15 eV . We see that this really is a method for qualitative understanding rather than quantitative results.

Finally, we will easily demonstrate the nature of chemical bonding using Hückel's model. If we take a molecule of ethylene (ethene), we are able to use Hückel's method to create a simple Hamiltonian. The matrix has the form

$$\mathbb{H} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}.$$

If we diagonalize this Hamiltonian, we get two states, one with energy $\alpha + \beta$ and the second with energy $\alpha - \beta$. Naturally, we place the two electrons in the former, since it has lower energy. Thus, the result yields a total energy of $2\alpha + 2\beta$. If we move the two carbon atoms to such a distance that the interaction between them is negligible, the β parameter describing the interaction is zero, while the α parameter is unchanged. In this case, both states have the same energy α . Thus, the interaction of the molecules "produces" a state with energy reduced by β . By occupying this state (orbital) we get an energy that is lower than if the two atoms involved were far apart. This orbital is the one keeping the atoms close together – it creates a chemical bond. It is therefore called a bonding orbital. On the other hand, the second orbital has higher energy than the isolated atoms when occupied; this one "sabotages" the chemical bond. It is therefore referred to as the antibonding orbital. In our model, the energy of the bonding orbital is lower by the same amount that the antibonding orbital is higher, but in reality, if we did not neglect the overlap between the orbitals, the antibonding would be higher by a little more than the bonding is lower.

So far, we have only described the coupling between p orbitals pointing perpendicular to the nuclear junction, which is only one of a pair of bonds in a double bond. The other is formed by p orbitals in the direction of the bond and can be found even in the ethane molecule. For this one, we could construct the same (approximate) Hamiltonian, just with a different parameter β . The parameter α could remain the same in this case, while this bond in ethylene is also formed by p orbitals, just rotated differently. This also gives us bonding and antibonding orbital for this bond.

In addition, there is a standard notation for the bonding and antibonding orbitals. For the bonding orbital, which is in the bond axis, the designation σ is used. The antibonding one is denoted by σ^* . Similarly, orbitals forming additional bonds in multiple bonds are denoted as π and π^* . In most molecules, the σ orbital lies below π , and conversely the σ^* orbital lies above π^* . If we have a triple bond, such as in the acetylene molecule, we have one occupied σ and two degenerate π bonding orbitals.

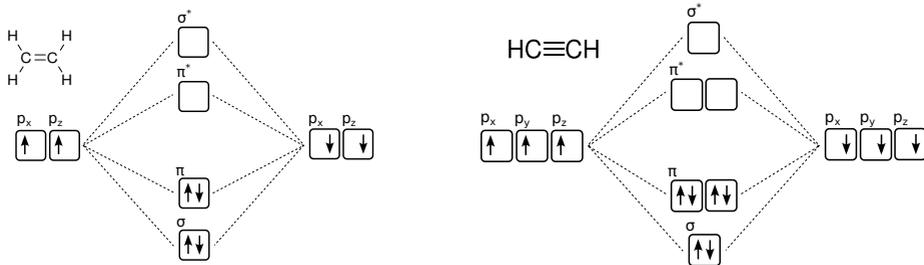


Fig. 2: Bonding and antibonding orbitals in the ethylene (left) and acetylene (right)

With these three episodes of the series, we've had some practice with calculating, and from now on we'll let the computer do all the counting for us.

FYKOS is organized by students of Faculty of Mathematics and Physics of Charles University. It's part of Media Communications and PR Office and is supported by Institute of Theoretical Physics of MFF UK, his employees and The Union of Czech Mathematicians and Physicists. The realization of this project was supported by Ministry of Education, Youth and Sports.

This work is licensed under Creative Commons Attribution-Share Alike 3.0 Unported. To view a copy of the license, visit <https://creativecommons.org/licenses/by-sa/3.0/>.