APPENDIX J

Orthogonalization

Schmidt Orthogonalization

Two Vectors

Imagine two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, each of length 1 (i.e., normalized), having the scalar product $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = a$. If $a = 0$, then the two vectors are orthogonal. We are interested in the case $a \neq 0$. Can one make such linear combinations of $\boldsymbol{u}$ and $\boldsymbol{v}$, that the new vectors $\boldsymbol{u}'$ and $\boldsymbol{v}'$ will be orthogonal? We can do that in many ways; two of them are called the *Schmidt orthogonalization:*
Case I:

$$\boldsymbol{u}' = \boldsymbol{u},$$
$$\boldsymbol{v}' = \boldsymbol{v} - \boldsymbol{u}\,\langle \boldsymbol{u}|\boldsymbol{v}\rangle\,,$$

Case II:

$$\boldsymbol{u}' = \boldsymbol{u} - \boldsymbol{v}\,\langle \boldsymbol{v}|\boldsymbol{u}\rangle$$
$$\boldsymbol{v}' = \boldsymbol{v}.$$

It is seen that the Schmidt orthogonalization is based on a very simple idea. In Case I, the first vector is left unchanged, while, one cuts out its component along the first one (Fig. J.1). In this way, the two vectors are treated differently (hence, the two cases above).

In this book, the vectors to orthogonalize will be the Hilbert space vectors (see Appendix B available at booksite.elsevier.com/978-0-444-59436-5); i.e., the normalized wave functions. In the case of two such vectors ϕ_1 and ϕ_2 having the scalar product $\langle\phi_1|\phi_2\rangle$, we construct the new orthogonal wave functions $\psi_1 = \phi_1, \psi_2 = \phi_2 - \phi_1\,\langle\phi_1|\phi_2\rangle$, or $\psi_1 = \phi_1 - \phi_2\,\langle\phi_2|\phi_1\rangle$, $\psi_2 = \phi_2$, analogous to the previous formulas.

More Vectors

In the case of many vectors, the procedure is similar. First, we decide about the order of the vectors to be orthogonalized. Then, we begin the procedure by leaving the first vector unchanged. Then, we continue, remembering that from a new vector, we have to cut out all its components

Ideas of Quantum Chemistry, Second Edition. http://dx.doi.org/10.1016/B978-0-444-59436-5.00030-1

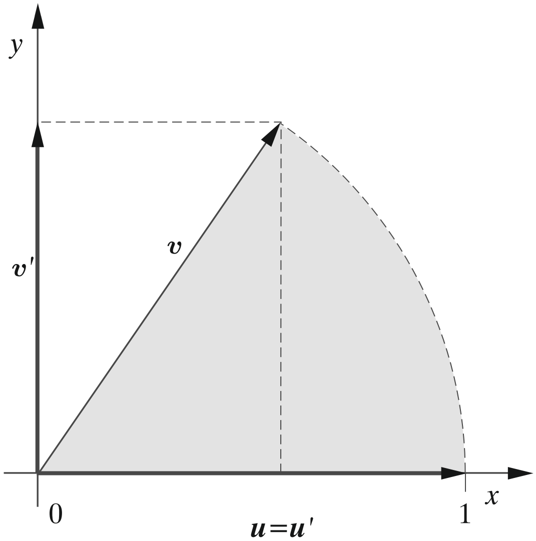

Fig. J.1. The Schmidt orthogonalization of the unit (i.e., normalized) vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. The new vectors are $\boldsymbol{u}'$ and $\boldsymbol{v}'$. Vector $\boldsymbol{u}' \equiv \boldsymbol{u}$, while from vector $\boldsymbol{v}$, we subtract its component along $\boldsymbol{u}$. The new vectors are orthogonal.

along the new vectors already found. Of course, the final set of vectors depends on the order chosen.

Löwdin Symmetric Orthogonalization

Imagine the normalized but non-orthogonal basis set wave functions collected as the components of the vector $\boldsymbol{\phi}$. By making proper linear combinations of the wave functions, we will get the orthogonal wave functions. The *symmetric* orthogonalization (as opposed to the Schmidt orthogonalization) treats all the wave functions on an equal footing. Instead of the old non-orthogonal basis set $\boldsymbol{\phi}$, we construct a new basis set $\boldsymbol{\phi}'$ by a linear transformation $\boldsymbol{\phi}' = \boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\phi}$, where $\boldsymbol{S}$ is the overlap matrix with the elements $S_{ij} = \langle \phi_i | \phi_j \rangle$, and the square matrix $\boldsymbol{S}^{-\frac{1}{2}}$ and its cousin $\boldsymbol{S}^{\frac{1}{2}}$ are defined in the following way. First, we diagonalize $\boldsymbol{S}$ using a unitary matrix $\boldsymbol{U}$; i.e., $\boldsymbol{U}^\dagger \boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^\dagger = \mathbf{1}$ (for real $\boldsymbol{S}$ the matrix $\boldsymbol{U}$ is orthogonal, $\boldsymbol{U}^T \boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^T = \mathbf{1}$):

$$\boldsymbol{S}_{diag} = \boldsymbol{U}^\dagger \boldsymbol{S} \boldsymbol{U}.$$

The eigenvalues of $\boldsymbol{S}$ are always positive, so the diagonal elements of $\boldsymbol{S}_{diag}$ can be replaced by their square roots, thus producing the matrix denoted by *symbol* $\boldsymbol{S}^{\frac{1}{2}}_{diag}$. Using this matrix, we define the matrices $\boldsymbol{S}^{\frac{1}{2}} = \boldsymbol{U}\boldsymbol{S}^{\frac{1}{2}}_{diag}\boldsymbol{U}^\dagger$ and $\boldsymbol{S}^{-\frac{1}{2}} = \left(\boldsymbol{S}^{\frac{1}{2}}\right)^{-1} = \boldsymbol{U}\boldsymbol{S}^{-\frac{1}{2}}_{diag}\boldsymbol{U}^\dagger$. Their *symbols* correspond to their properties: $\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{S}^{\frac{1}{2}} = \boldsymbol{U}S_{diag}^{\frac{1}{2}}\boldsymbol{U}^{\dagger}\boldsymbol{U}S_{diag}^{\frac{1}{2}}\boldsymbol{U}^{\dagger} = \boldsymbol{U}S_{diag}^{\frac{1}{2}}\boldsymbol{S}_{diag}^{\frac{1}{2}}\boldsymbol{U}^{\dagger} = \boldsymbol{U}S_{diag}\boldsymbol{U}^{\dagger} = \boldsymbol{S}$; similarly, $\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{S}^{-\frac{1}{2}} = \boldsymbol{S}^{-1}$. Also, a straightforward calculation gives[1] $\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{S}^{\frac{1}{2}} = \mathbf{1}$.

An important feature of the symmetric orthogonalization is[2] that among all possible orthogonalizations, the symmetric orthogonalization ensures

$$\sum_i \left|\left|\phi_i - \phi_i'\right|\right|^2 = \text{minimum},$$

where $\left|\left|\phi_i - \phi_i'\right|\right|^2 \equiv \left\langle \phi_i - \phi_i' | \phi_i - \phi_i' \right\rangle$. This means that

> the symmetrically orthogonalized functions ϕ_i' are the least distant in the Hilbert space from the original functions ϕ_i. Thus, a symmetric orthogonalization indicates the gentlest pushing of the directions of the vectors in order to get them orthogonal.

Example

The symmetric orthogonalization will be shown taking an example of two non-orthogonal vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ (instead of functions ϕ_1 and ϕ_2), each of length 1 with the scalar product[3] $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = a \neq 0$. We decide to consider the vectors with real components; hence $a \in R$. First, we have to construct matrix $\boldsymbol{S}^{-\frac{1}{2}}$. Here is how we arrive at this. Matrix $\boldsymbol{S}$ is equal $\boldsymbol{S} = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$, and as we see it is symmetric. First, let us diagonalize $\boldsymbol{S}$. To achieve that, we apply the orthogonal transformation $\boldsymbol{U}^{\dagger}\boldsymbol{S}\boldsymbol{U}$ (thus, in this case $\boldsymbol{U}^{\dagger} = \boldsymbol{U}^{T}$), where (to assure the orthogonality of the transformation matrix) we choose $\boldsymbol{U} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, and therefore $\boldsymbol{U}^{\dagger} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, with angle θ to be specified. After the transformation, we have: $\boldsymbol{U}^{\dagger}\boldsymbol{S}\boldsymbol{U} = \begin{pmatrix} 1 - a\sin 2\theta & a\cos 2\theta \\ a\cos 2\theta & 1 + a\sin 2\theta \end{pmatrix}$.

---

1 The matrix $\boldsymbol{S}^{-\frac{1}{2}}$ is not just a symbol anymore. Let us check whether the transformation $\boldsymbol{\phi}' = \boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\phi}$ gives orthonormal wave functions (vectors). Remembering that $\boldsymbol{\phi}$ represents a vertical vector with the components ϕ_i (being functions): $\int \boldsymbol{\phi}\boldsymbol{\phi}^T d\tau = \boldsymbol{S}$, while $\int \boldsymbol{\phi}'\boldsymbol{\phi}'^T d\tau = \int \boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\phi}\boldsymbol{\phi}^T\boldsymbol{S}^{-\frac{1}{2}} d\tau = \boldsymbol{S}^{-1}\int \boldsymbol{\phi}\boldsymbol{\phi}^T d\tau = \boldsymbol{S}^{-1}\boldsymbol{S} = \mathbf{1}$. This is what we wanted to show.

2 G.W. Pratt and S.P. Neustadter, *Phys.Rev.*, *101*, 1248 (1956).

3 $-1 \leq a \leq 1$.

It is seen that if we chose $\theta = 45^o$, then the matrix $\boldsymbol{U}^\dagger \boldsymbol{S} \boldsymbol{U}$ would be *diagonal*[4] (that is what we would like to have): $\boldsymbol{S}_{diag} = \begin{pmatrix} 1-a & 0 \\ 0 & 1+a \end{pmatrix}$. We construct, then, $\boldsymbol{S}_{diag}^{\frac{1}{2}} = \begin{pmatrix} \sqrt{1-a} & 0 \\ 0 & \sqrt{1+a} \end{pmatrix}$.

Next, we form $\boldsymbol{S}^{\frac{1}{2}} = \boldsymbol{U}\boldsymbol{S}_{diag}^{\frac{1}{2}}\boldsymbol{U}^\dagger = \frac{1}{2}\begin{pmatrix} \sqrt{1-a}+\sqrt{1+a} & \sqrt{1+a}-\sqrt{1-a} \\ \sqrt{1+a}-\sqrt{1-a} & \sqrt{1-a}+\sqrt{1+a} \end{pmatrix}$ and the matrix $\boldsymbol{S}^{-\frac{1}{2}}$ needed[5] for the transformation is equal to

$$\boldsymbol{S}^{-\frac{1}{2}} = \boldsymbol{U}\boldsymbol{S}_{diag}^{-\frac{1}{2}}\boldsymbol{U}^\dagger = \boldsymbol{U}\begin{pmatrix} \frac{1}{\sqrt{1-a}} & 0 \\ 0 & \frac{1}{\sqrt{1+a}} \end{pmatrix}\boldsymbol{U}^\dagger = \frac{1}{2}\begin{pmatrix} \frac{1}{\sqrt{1-a}}+\frac{1}{\sqrt{1+a}} & \frac{1}{\sqrt{1+a}}-\frac{1}{\sqrt{1-a}} \\ \frac{1}{\sqrt{1+a}}-\frac{1}{\sqrt{1-a}} & \frac{1}{\sqrt{1-a}}+\frac{1}{\sqrt{1+a}} \end{pmatrix}.$$

Now we are ready to construct the orthogonalized vectors[6]:

$$\begin{pmatrix} \boldsymbol{u}' \\ \boldsymbol{v}' \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \frac{1}{\sqrt{1-a}}+\frac{1}{\sqrt{1+a}} & \frac{1}{\sqrt{1+a}}-\frac{1}{\sqrt{1-a}} \\ \frac{1}{\sqrt{1+a}}-\frac{1}{\sqrt{1-a}} & \frac{1}{\sqrt{1-a}}+\frac{1}{\sqrt{1+a}} \end{pmatrix}\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix}.$$

$$\boldsymbol{u}' = C\boldsymbol{u} + c\boldsymbol{v}$$
$$\boldsymbol{v}' = c\boldsymbol{u} + C\boldsymbol{v},$$

where the "*large*" coefficient $C = \frac{1}{2}\left(\frac{1}{\sqrt{1-a}}+\frac{1}{\sqrt{1+a}}\right)$, and a "*small*" admixture $c = \frac{1}{2}\left(\frac{1}{\sqrt{1+a}}-\frac{1}{\sqrt{1-a}}\right)$. As we can see, the new (orthogonal) vectors are formed from the old ones (non-orthogonal) by an *identical* (hence the name *symmetric orthogonalization*) admixture of the old vectors; i.e., the contribution of $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\boldsymbol{u}'$ is the same as that of $\boldsymbol{v}$ and $\boldsymbol{u}$ in $\boldsymbol{v}'$.

The new vectors are obtained by correcting the directions of the old ones, each by the same angle.

This is illustrated in Fig. J.2.

The new vectors automatically have a length of 1, the same as the starting vectors.

4 In such a case, the transformation matrix is

$$\boldsymbol{U} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

5 They are symmetric matrices. For example, $\left(\boldsymbol{S}^{\frac{1}{2}}\right)_{ij} = \left(\boldsymbol{U}\boldsymbol{S}_{diag}^{\frac{1}{2}}\boldsymbol{U}^\dagger\right)_{ij} = \sum_k \sum_l U_{ik}\left(\boldsymbol{S}_{diag}^{\frac{1}{2}}\right)_{kl} U_{jl} = \sum_k \sum_l U_{ik}\left(\boldsymbol{S}_{diag}^{\frac{1}{2}}\right)_{kl} \delta_{kl} U_{jl} = \sum_k U_{ik}\left(\boldsymbol{S}_{diag}^{\frac{1}{2}}\right)_{kk} U_{jk} = \left(\boldsymbol{S}^{\frac{1}{2}}\right)_{ji}$.

6 It is seen that if the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ were already orthogonal (i.e., $a = 0$), then $\boldsymbol{u}' = \boldsymbol{u}$ and $\boldsymbol{v}' = \boldsymbol{v}$. Of course, we like this result.

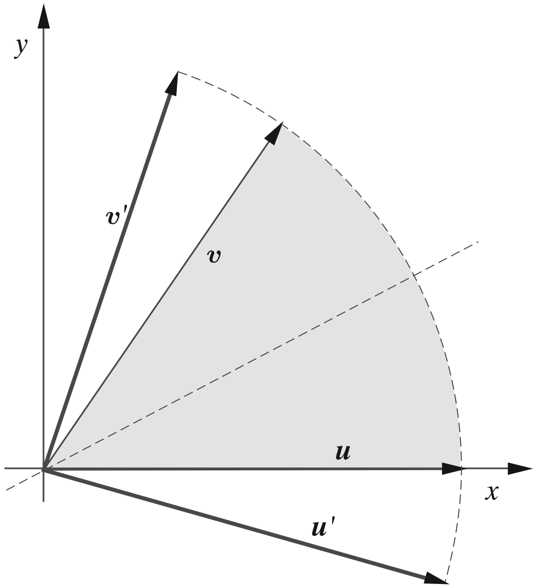

Fig. J.2. The symmetric (or Löwdin's) orthogonalization of the normalized vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. The vectors are pushed off by the same angle in such a way as to assure $\boldsymbol{u}'$ and $\boldsymbol{v}'$ to become orthogonal.