J. ORTHOGONALIZATION

1 SCHMIDT ORTHOGONALIZATION

Two vectors

Imagine two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, each of length 1 (i.e. normalized), with the dot product $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = a$. If $a = 0$, the two vectors are orthogonal. We are interested in the case $a \neq 0$. Can we make such linear combinations of $\boldsymbol{u}$ and $\boldsymbol{v}$, so that the new vectors, $\boldsymbol{u}'$ and $\boldsymbol{v}'$, will be orthogonal? We can do this 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 Schmidt orthogonalization is based on a very simple idea. In Case I the first vector is left unchanged, while from the second vector, we cut out its component along the first (Fig. J.1). In this way the two vectors are treated differently (hence, the two cases above).

In this book the vectors we orthogonalize will be Hilbert space vectors (see Appendix B), i.e. the normalized wave functions. In the case of two such vectors ϕ_1 and ϕ_2 having a dot product $\langle \phi_1|\phi_2\rangle$ we construct the new orthogonal wave

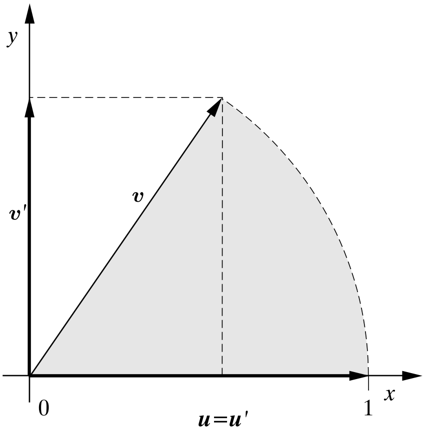

Fig. J.1. The Schmidt orthogonalization of the unit (i.e. normalized) vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. The new vectors are $\boldsymbol{u}'$ and $\boldsymbol{v}'$. Vector $\boldsymbol{u}' \equiv \boldsymbol{u}$, while from vector $\boldsymbol{v}$ we subtract its component along $\boldsymbol{u}$.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

functions $\psi_1 = \phi_1$, $\psi_2 = \phi_2 - \phi_1\langle\phi_1|\phi_2\rangle$ or $\psi_1 = \phi_1 - \phi_2\langle\phi_2|\phi_1\rangle$, $\psi_2 = \phi_2$ by analogy to the previous formulae.

More vectors

In case of many vectors the procedure is similar. First, we decide 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 along the new vectors already found. Of course, the final set of vectors depends on the order chosen.

2 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, $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^T = \mathbf{1}$),

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

The eigenvalues of $\boldsymbol{S}$ are always positive, therefore the diagonal elements of $\boldsymbol{S}_{\mathrm{diag}}$ can be replaced by their square roots, thus producing the matrix denoted by the *symbol* $\boldsymbol{S}^{\frac{1}{2}}_{\mathrm{diag}}$. Using the latter matrix we define the matrices

$$\boldsymbol{S}^{\frac{1}{2}} = \boldsymbol{U}\boldsymbol{S}^{\frac{1}{2}}_{\mathrm{diag}}\boldsymbol{U}^\dagger \quad \text{and} \quad \boldsymbol{S}^{-\frac{1}{2}} = \left(\boldsymbol{S}^{\frac{1}{2}}\right)^{-1} = \boldsymbol{U}\boldsymbol{S}^{-\frac{1}{2}}_{\mathrm{diag}}\boldsymbol{U}^\dagger.$$

Their *symbols* correspond to their properties:

$$\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{S}^{\frac{1}{2}} = \boldsymbol{U}\boldsymbol{S}^{\frac{1}{2}}_{\mathrm{diag}}\boldsymbol{U}^\dagger\boldsymbol{U}\boldsymbol{S}^{\frac{1}{2}}_{\mathrm{diag}}\boldsymbol{U}^\dagger = \boldsymbol{U}\boldsymbol{S}^{\frac{1}{2}}_{\mathrm{diag}}\boldsymbol{S}^{\frac{1}{2}}_{\mathrm{diag}}\boldsymbol{U}^\dagger = \boldsymbol{U}\boldsymbol{S}_{\mathrm{diag}}\boldsymbol{U}^\dagger = \boldsymbol{S},$$

similarly $\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{S}^{-\frac{1}{2}} = \boldsymbol{S}^{-1}$. Also, a straightforward calculation gives[1] $\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{S}^{\frac{1}{2}} = \mathbf{1}$.

---

[1]The matrix $\boldsymbol{S}^{-\frac{1}{2}}$ is no longer a symbol anymore. Let us check whether the transformation $\boldsymbol{\phi}' = \boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\phi}$ indeed gives orthonormal wave functions (vectors). Remembering that $\boldsymbol{\phi}$ represents a vertical vector with components ϕ_i (being functions): $\int \boldsymbol{\phi}^*\boldsymbol{\phi}^T\,\mathrm{d}\tau = \boldsymbol{S}$, while $\int \boldsymbol{\phi}'^*\boldsymbol{\phi}'^T\,\mathrm{d}\tau = \int \boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\phi}^*\boldsymbol{\phi}^T\boldsymbol{S}^{-\frac{1}{2}}\,\mathrm{d}\tau = \mathbf{1}$. This is what we wanted to show.

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

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

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

> the symmetrically orthogonalized functions ϕ_i' are the "least distant" from the original functions ϕ_i. Thus symmetric orthogonalization means a gentle pushing the directions of the vectors in order to get them to be orthogonal.

Example

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

$$\boldsymbol{U} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad \text{and therefore} \quad \boldsymbol{U}^\dagger = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

with angle θ to be specified. After the transformation we have:

$$\boldsymbol{U}^\dagger \boldsymbol{S} \boldsymbol{U} = \begin{pmatrix} 1 - a\sin 2\theta & a\cos 2\theta \\ a\cos 2\theta & 1 + a\sin 2\theta \end{pmatrix}.$$

We see that if we chose $\theta = 45°$, the matrix $\boldsymbol{U}^\dagger \boldsymbol{S} \boldsymbol{U}$ will be *diagonal*[4] (this is what we would like to have):

$$\boldsymbol{S}_{\text{diag}} = \begin{pmatrix} 1 - a & 0 \\ 0 & 1 + a \end{pmatrix}.$$

We then construct

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

---

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

[3] $-1 \leqslant a \leqslant 1$.

[4] In such a case the transformation matrix is

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

Next, we form[5]

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

and the matrix $S^{-\frac{1}{2}}$ needed for the transformation is equal to

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

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

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

$$u' = Cu + cv,$$

$$v' = cu + Cv.$$

where the "large" coefficient

$$C = \frac{1}{2}\left(\frac{1}{\sqrt{1-a}} + \frac{1}{\sqrt{1+a}}\right),$$

and there is 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 u and v in u' is the same as that of v and u in 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.

[5]They are symmetric matrices. For example,

$$(S^{\frac{1}{2}})_{ij} = (US^{\frac{1}{2}}_{\text{diag}}U^{\dagger})_{ij} = \sum_k\sum_l U_{ik}(S^{\frac{1}{2}}_{\text{diag}})_{kl}U_{jl} = \sum_k\sum_l U_{ik}(S^{\frac{1}{2}}_{\text{diag}})_{kl}\delta_{kl}U_{jl}$$
$$= \sum_k U_{ik}(S^{\frac{1}{2}}_{\text{diag}})_{kk}U_{jk} = (S^{\frac{1}{2}})_{ji}.$$

[6]We see that if the vectors u and v were already orthogonal, i.e. $a = 0$, then $u' = u$ and $v' = v$. Of course, we like this result.

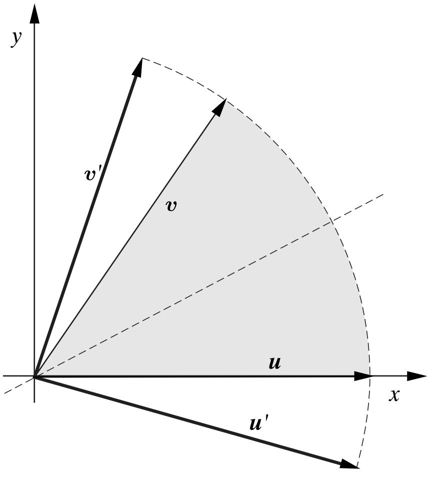

Fig. J.2. The symmetric (or Löwdin's) orthogonalization of the normalized vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. The vectors are just pushed away by the same angle in such a way as to ensure $\boldsymbol{u}'$ and $\boldsymbol{v}'$ become orthogonal.

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