Revista Mexicana de Física 31 No. 4 (1985) 743-758

ORTHONORMALIZATION ON THE PLANE: A GEOMETRIC APPROACH

C.E. Soliverez[+]

Centro Atómico Bariloche[‡] and Instituto Balseiro[#]

Casilla de Correo 439, 8400 Bariloche, Argentina

and

E.R. Gagliano[+]

Instituto de Desarrollo Tecnológico para la Industria Química,

Casilla de Correo 91, 3000 Santa Fé, Argentina

(recibido junio 4, 1984; aceptado marzo 8, 1985)

ABSTRACT

When confining oneself to the two-dimensional case it is possible to give simple geometric interpretations to the properties of alternative orthonormalization schemes. The symmetry, proximity and localization properties of the symmetric orthonormalization, and the delocalization property of the canonical scheme, among others, become thus evident at first sight.

[+] Affiliated to CONICET.

[‡] Comisión Nacional de Energía Atómica.

[#] Comisión Nacional de Energía Atómica and Universidad Nacional de Cuyo.

RESUMEN

Al restringirse al caso bidimensional es posible dar interpretaciones geométricas sencillas de los diferentes esquemas de ortonormalización utilizados en física y química. Las propiedades de simetría, proximidad y localización de la ortonormalización simétrica, y la de delocalización del esquema canónico, entre otras, resultan entonces evidentes a primera vista.

1. INTRODUCTION

Orthonormal bases are cuestomarily used when solving physical and chemical problems, because then the mathematical formulation usually becomes simpler. There are nevertheless situations where the natural bases are nonorthonormal. That is for instance the case of the basis vectors of crystallography, or the atomic orbitals used in molecular and solid-state calculations. This is so in the last case because different orbitals, although orthogonal when belonging to the same atom, have nonvanishing overlap when centered on different atoms.

As textbooks on mathematics discuss solely the Gram-Schmidt procedure[1] one may get the impression that this is the only known orthonormalization scheme. There are in fact infinitely many such schemes, the Gram-Schmidt one being probably the most cumbersome of all.

While the properties of all alternative orthonormalization schemes are long since well known[2], the necessity of working in a many-dimensional vectorial space precludes an easy visualization of the mathematical results. If one starts by considering the two-dimensional case the orthonormalization problem becomes instead a simple geometrical problem on the plane which may be solved almost by inspection. It is then easy to obtain a good grasp of the many important differences between the available procedures. As the two-dimensional case has an interest of its own (remember molecular orbitals) and the better understanding thus obtained is valuable for the general many-dimensional case, a detailed discussion is well warranted.

2. ORTHONORMALIZATION ON THE PLANE

Let ϕ_1 and ϕ_2 be two real, nonorthogonal, linearly independent, and normalized wavefunctions:

$$\begin{aligned} &\langle\phi_1|\phi_1\rangle = \langle\phi_2|\phi_2\rangle = 1 \quad , \\ &\langle\phi_1|\phi_2\rangle = \langle\phi_2|\phi_1\rangle = S \quad , \end{aligned} \tag{1}$$

where S is the overlap. If we take them to be vectors on the plane we may write instead

$$\langle\phi_j|\phi_k\rangle = g_{jk} = \underset{\sim}{\phi}_j \cdot \underset{\sim}{\phi}_k \quad , \tag{2}$$

g_{jk} being an element of the symmetric metric matrix

$$\underset{\sim}{g} = \begin{pmatrix} 1 & S \\ S & 1 \end{pmatrix} \quad . \tag{3}$$

Linear independence implies that $\underset{\sim}{\phi}_1$ and $\underset{\sim}{\phi}_2$ are not parallel, that is (see Fig. 1)

$$\begin{aligned} &-1 < S = \underset{\sim}{\phi}_1 \cdot \underset{\sim}{\phi}_2 = \cos\gamma < 1 \quad , \\ &\text{Det } \underset{\sim}{g} = 1 - S^2 \neq 0 \quad , \end{aligned} \tag{4}$$

thus showing $\underset{\sim}{g}$ to be nonsingular.

Our problem is finding a real transformation $\underset{\sim}{0}$ such that the new vectors

$$\underset{\sim}{\phi}'_j = \sum_{\ell} 0_{\ell j} \, \underset{\sim}{\phi}_\ell \tag{5}$$

are orthonormal. That is

$$\underset{\sim}{\phi}'_j \cdot \underset{\sim}{\phi}'_k = \delta_{jk} \quad , \qquad \underset{\sim}{0}^t \cdot \underset{\sim}{g} \cdot \underset{\sim}{0} = \underset{\sim}{1} \quad , \tag{6}$$

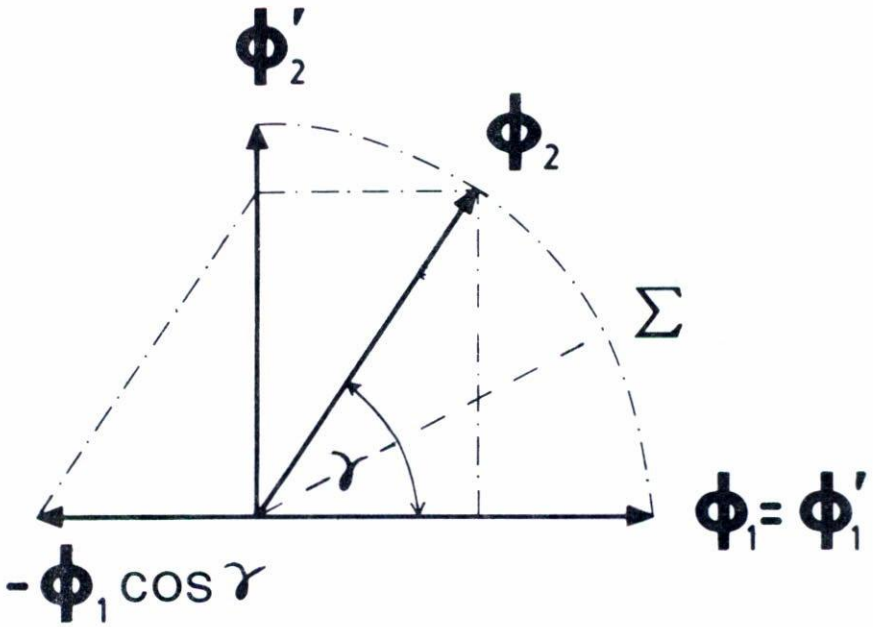

Fig. 1. Two unit vectors ϕ_1 and ϕ_2 are linearly independent if the angle γ between them is different from 0, π. A vector ϕ_2' normal to ϕ_1 may then be obtained by the Gram-Schmidt procedure Eq. (8). The deshed line $\sum$ is a mirror line.

where $\underset{\sim}{0}^t$ is the transpose of $\underset{\sim}{0}$, and $\underset{\sim}{1}$ the unit matrix. The general solution of Eq. (6) is

$$\underset{\sim}{0} = \underset{\sim}{g}^{-\frac{1}{2}} \cdot \underset{\sim}{U} \quad , \qquad \underset{\sim}{U}^t = \underset{\sim}{U}^{-1} \quad , \tag{7}$$

where $\underset{\sim}{U}$ is an arbitrary orthogonal matrix. In the general case of complex vectors the hermitian adjoint should be taken instead of the transpose, and $\underset{\sim}{U}$ should be unitary.

The best known orthonormalization method is the Gram-Schmidt process[1] where starting with a given vector $\underset{\sim}{\phi}_1$ is a normal vector $\underset{\sim}{\phi}_2'$ is obtained from

$$\underset{\sim}{\phi}_2' = (\underset{\sim}{\phi}_2 - \cos\gamma\ \underset{\sim}{\phi}_1)\Big/\|\underset{\sim}{\phi}_2 - \cos\gamma\ \underset{\sim}{\phi}_1\| \quad . \tag{8}$$

It is quite clear that the Gram-Schmidt process gives the leading role to the initial vector thus destroying the symmetry properties of the primitive set. Assume for instance that $\underset{\sim}{\phi}_1$ and $\underset{\sim}{\phi}_2$ are related by the mirror line $\sum$ of Fig. 1,

$$\underset{\sim}{\sum} \cdot \underset{\sim}{\phi}_1 = \underset{\sim}{\phi}_2 \quad , \qquad \underset{\sim}{\sum} \cdot \underset{\sim}{\phi}_2 = \underset{\sim}{\phi}_1 \quad , \tag{9}$$

where $\underset{\sim}{\sum}$ is the diadic performing the reflection. It is immediately seen that the reflection symmetry is not preserved for the new basis, this being a consequence of the unequal weight given to each primitive vector in the construction of the respective new one, as measured by the projection $\underset{\sim}{\phi}_j \cdot \underset{\sim}{\phi}_j'$. An alternative procedure is suggested at once where the new orthogonal vectors are symmetrically rotated respect to the old ones, as shown in Fig. 2. It may be seen that the reflection symmetry is now preserved for the new vectors. As a matter of fact all unitary relationships between basis vectors are preserved in this orthonormalization process[3] which corresponds to Löwdin's symmetric orthonormalization

$$\underset{\sim}{0} = \underset{\sim}{g}^{-\frac{1}{2}} \quad . \tag{10}$$

We will now find $\underset{\sim}{g}^{-\frac{1}{2}}$. It should first be notices that any square root of $\underset{\sim}{g}$ (there are four of them) must verify

$$\underset{\sim}{g}^{\frac{1}{2}} = \begin{pmatrix} p & q \\ q & p \end{pmatrix} \quad , \qquad \underset{\sim}{g}^{\frac{1}{2}} \cdot \underset{\sim}{g}^{\frac{1}{2}} = \underset{\sim}{g} \quad , \tag{11}$$

from which we obtain

$$\begin{aligned} p^2 + q^2 &= 1 \quad , & 2pq &= \cos\gamma \quad , \\ p &= \cos\alpha \quad , & q &= \sin\alpha \quad , \end{aligned} \tag{12}$$

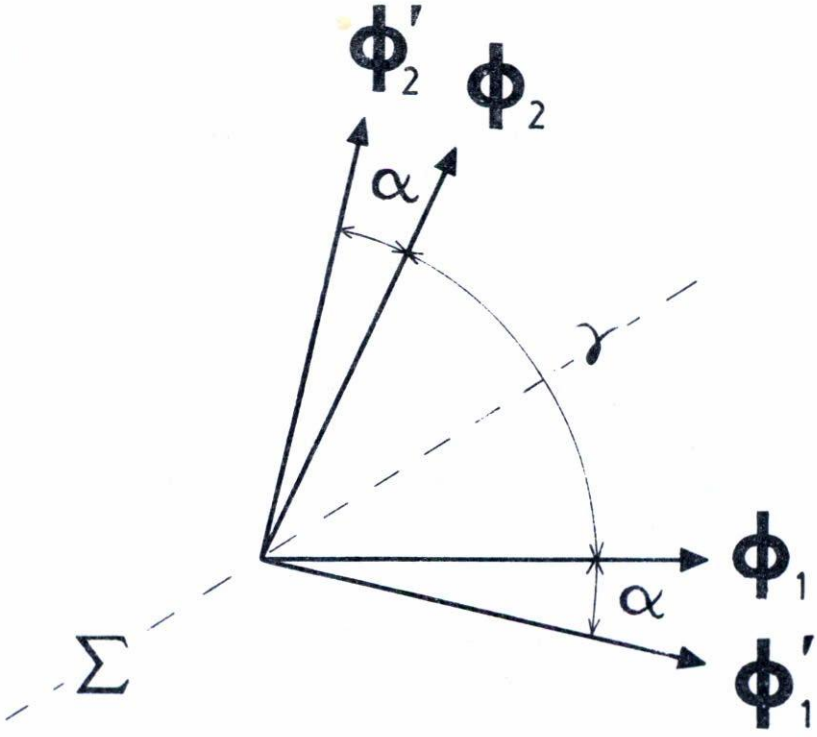

Fig. 2. The symmetric orthonormalization process generates new vectors ϕ_1', ϕ_2' which are symmetrically rotated respect to the original ones, thus preserving orthogonal relationships such as the mirror line $\sum$.

$$\sin 2\alpha = \cos \gamma \quad . \tag{13}$$

Because $\gamma \neq 0, \pi$ it is easily seen from Fig. 3 that Eq. (13) always has four different solutions in the interval $(-\pi, \pi)$, as given by

$$\alpha = \frac{\pi}{4} \pm \frac{\gamma}{2} + n\pi \quad , \tag{14}$$

where the choice of the integer n depends on γ. The inverse matrix $\underset{\sim}{g}^{-\frac{1}{2}}$ may now be easily written

$$\underset{\sim}{g}^{-\frac{1}{2}} = \begin{pmatrix} \cos\alpha/\cos 2\alpha & -\sin\alpha/\cos 2\alpha \\ -\sin\alpha/\cos 2\alpha & \cos\alpha/\cos 2\alpha \end{pmatrix} . \qquad (15)$$

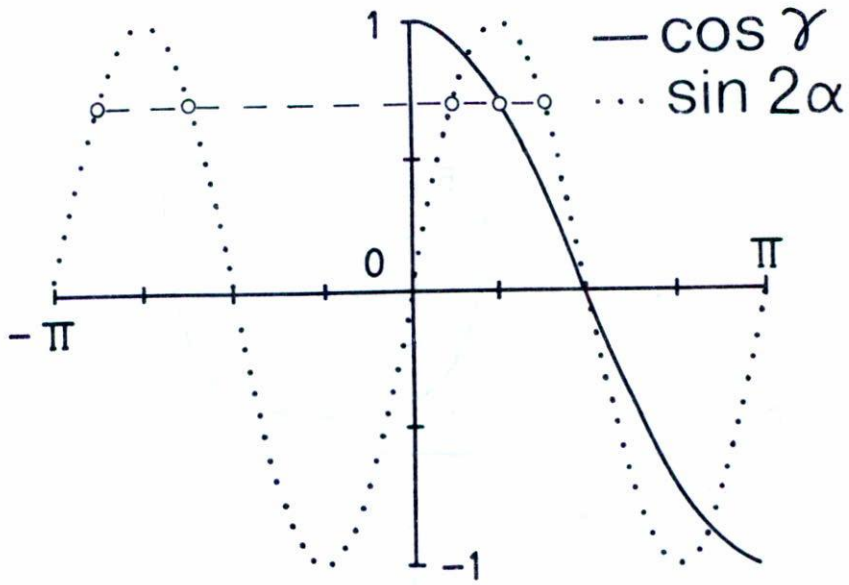

Fig. 3. For any given value of γ in $(0, \pi)$ $(\gamma \neq 0, \pi)$ there are always four different solutions to the equation $\sin 2\alpha = \cos\gamma$ which determines $\underset{\sim}{g}^{-\frac{1}{2}}$.

The four different values of α correspond to the four different combinations of signs of the eigenvalues of $\underset{\sim}{g}^{-\frac{1}{2}}$ which are obtained from the secular equation

$$\left(\frac{\cos\alpha}{\cos 2\alpha} - \lambda\right)^2 - \left(\frac{\sin\alpha}{\cos 2\alpha}\right)^2 = 0 \quad , \qquad (16)$$

namely

$$\lambda^{\pm} = (\cos\alpha \mp \sin\alpha)/\cos 2\alpha = \pm\, 1/\sqrt{2}\, \sin(\alpha \pm \pi/4), \qquad (17)$$

eigenvalues which are plotted in Fig. 4. Replacing (10) and (15) into (5) one obtains

$$\underset{\sim}{\phi}_1' = \frac{\cos\alpha}{\cos 2\alpha}\,\underset{\sim}{\phi}_1 - \frac{\sin\alpha}{\cos 2\alpha}\,\underset{\sim}{\phi}_2 \quad ,$$

$$\underset{\sim}{\phi}_2' = \frac{-\sin\alpha}{\cos 2\alpha}\,\underset{\sim}{\phi}_1 \quad \frac{\cos\alpha}{\cos 2\alpha}\,\underset{\sim}{\psi}_2 \tag{18}$$

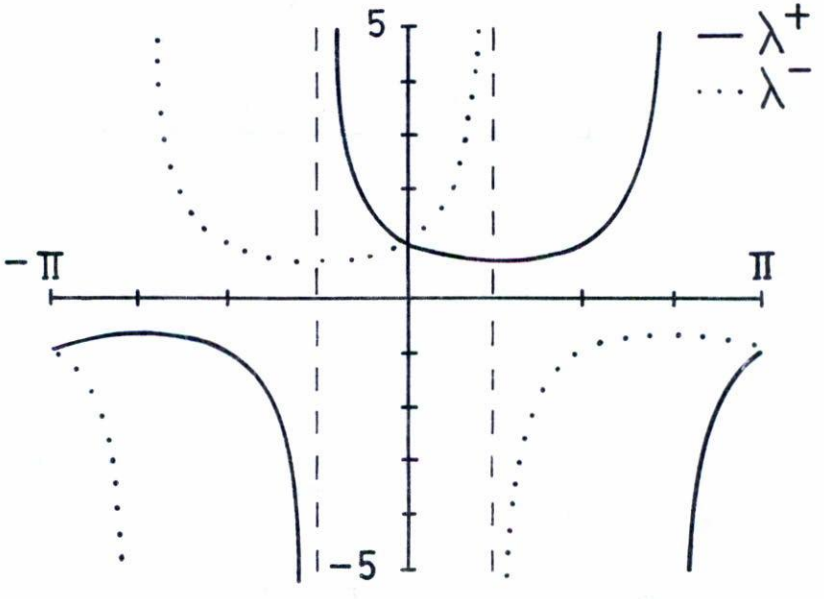

Fig. 4. Eingenvalues $\lambda^{\pm}$ of $g^{-\frac{1}{2}}$ as a function of α . A comparison with Fig. 3 shows the branch $-\pi/4 \leqslant \alpha \leqslant \pi/4$ to be the only one which makes $\underset{\sim}{g}^{-\frac{1}{2}}$ positive definite.

The new vectors are easily verified to be orthonormalized and to correspond to a rotation in α of the primitive vectors, as may be seen from

$$\underset{\sim}{\phi}_1 \cdot \underset{\sim}{\phi}_1' = \underset{\sim}{\phi}_2 \cdot \underset{\sim}{\phi}_2' = (\cos\alpha - \sin\alpha \cos\gamma)/\cos 2\alpha = \cos\alpha \ , \tag{19}$$

where use has been made of Eq. (13). One may see from Fig. 4 that both

eigenvalues are positive when

$$-\pi/4 \leq a \leq \pi/4, \qquad \alpha = \pi/4 - \gamma/2 \tag{20}$$

this being the only branck in Fig. 3 which makes $g^{-\frac{1}{2}}$ positive definite[4]. This solution is the one depicted in Fig. 2, and has the unique property of giving orthonormalized vectors which are the closest to the old ones in the sense that

$$\Delta^2 = \|\underset{\sim}{\phi}_1' - \underset{\sim}{\phi}_1\|^2 + \|\underset{\sim}{\phi}_2' - \underset{\sim}{\phi}_2\|^2 \tag{21}$$

is a minimum. This property, which holds for the general many-dimensional case[5], may be easily proved here. To that end consider as in Fig. 5 an arbitrary pair of orthonormal vectors, where

$$\|\underset{\sim}{\phi}_j' - \underset{\sim}{\phi}_j\| = |2 \sin(\alpha_j/2)|, \qquad \gamma = \alpha_1 + \alpha_2 + \pi/2 \quad . \tag{22}$$

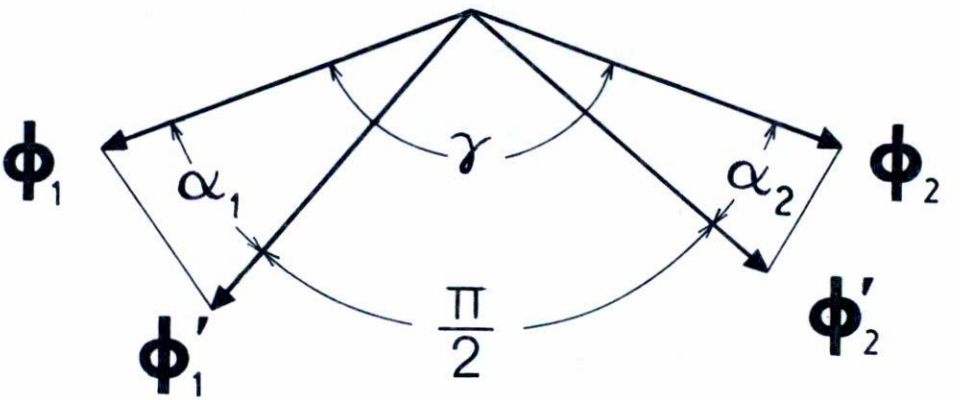

Fig. 5. For a general orthonormalized pair $\underset{\sim}{\phi}_1'$, $\underset{\sim}{\phi}_2'$ the angles $\underset{\sim}{\alpha}_1$ and $\underset{\sim}{\alpha}_2$ will be different. The case is here illustrated where $\gamma > \pi/2$ for which it is proved in the text that Δ^2 Eq. (21) is a minimum when $\alpha_1 = \alpha_2$.

Therefore

$$(\Delta/2)^2 = \sin^2(\alpha_1/2) + \sin^2(\alpha_2/2) = \sin^2(\alpha_1/2) + \sin^2[(\gamma - \alpha_1 - \pi/2)/2]$$

$$= 1 - \frac{1}{2}\cos\alpha_1 - \frac{1}{2}\cos(\gamma - \alpha_1 - \pi/2) \quad , \tag{23}$$

which is easily verified to be a minimum if

$$\alpha_1 = -\pi/4 + \gamma/2 = |\alpha_2| \quad . \tag{24}$$

A similar proof holds if $\gamma > \pi/2$. The other three solutions for a yield orthonormalized vectors which do not have the proximity property and are related to ϕ_1' and ϕ_2' Eq. (20) by

$$\begin{aligned} \phi_1'' &= -\phi_1' \quad , & \phi_2'' &= -\phi_2' \quad , \\ \phi_1''' &= \phi_2' \quad , & \phi_2''' &= \phi_1' \quad , \\ \phi_1'''' &= -\phi_2' \quad , & \phi_2'''' &= -\phi_1' \quad , \end{aligned} \tag{25}$$

In Fig. 6 we plot the values of the coeficients in Eq. (18) corresponding to the positive-definite branch Eq. (20). They are seen to blow-up in the linearly-dependent case $S = \cos\gamma = \pm 1$. The reason for this behaviour may be understood in the following fashion. Let us first define

$$\phi^{\pm} = \frac{1}{2}(\phi_1 \pm \phi_2) \quad , \tag{26}$$

from which it follows that

$$\begin{aligned} \phi_1' &= \phi^+/\sqrt{2}\cos(\gamma/2) + \phi^-/\sqrt{2}\sin(\gamma/2) \quad , \\ \phi_2' &= \phi^+/\sqrt{2}\cos(\gamma/2) - \phi^-/\sqrt{2}\sin(\gamma/2) \quad , \end{aligned} \tag{27}$$

thus showing that the symmetrically orthonormalized vectors may be obtained as a linear combination of the already normal vectors ϕ^+, ϕ^- as shown in Fig. 7. When ϕ_2 approaches $-\phi_1$ (that is $\cos\gamma = S$ approaches -1) ϕ^+

becomes smaller thus requiring a larger coefficient in order to make $\underset{\sim}{\phi}'_j$ turn further away from $\underset{\sim}{\phi}_j$. In the limit $\underset{\sim}{\phi}_2 = -\underset{\sim}{\phi}_1$ an infinitely large coefficient would be required. A similar analysis holds for the limit in which $\underset{\sim}{\phi}_2$ becomes parallel to $\underset{\sim}{\phi}_1$. From a mathematical point of view the divergence appears because one of the eigenvalues of $\underset{\sim}{g}$ vanishes thus making one of the eigenvalues of $\underset{\sim}{g}^{-\frac{1}{2}}$ to diverge.

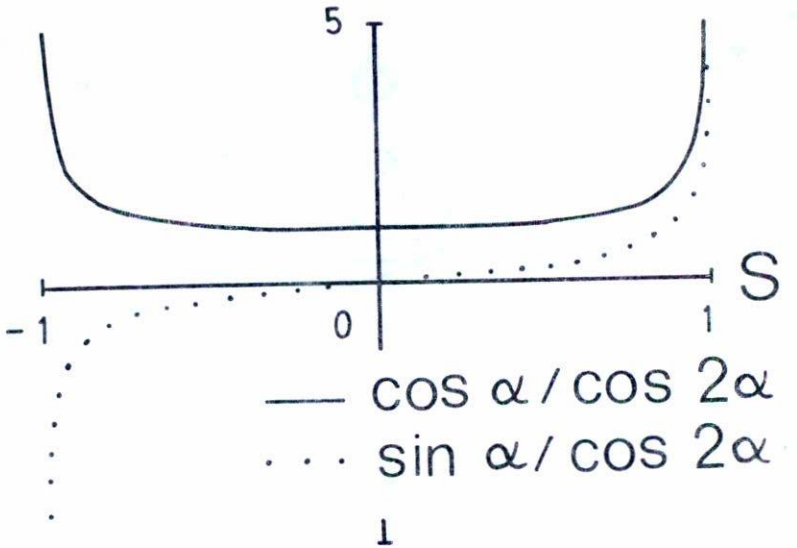

Fig. 6. Coefficients of the symmetrically orthonormalized vectors Eq. (18) as functions of the overlap S = cos γ. The coefficients diverge for the linearly dependent case S = 1.

As $\underset{\sim}{\phi}^+$ and $\underset{\sim}{\phi}^-$ are always orthogonal, one may wonder to which particular transformation $\underset{\sim}{0}$ they correspond. To that end we will first consider the matrix $\underset{\sim}{S}$ which diagonalizes $\underset{\sim}{g}^{-\frac{1}{2}}$. It may be easily verified that

$$\underset{\sim}{S}^t \cdot \underset{\sim}{g}^{-\frac{1}{2}} \cdot \underset{\sim}{S} = \begin{pmatrix} \lambda^+ & 0 \\ 0 & \lambda^- \end{pmatrix} , \qquad (28)$$

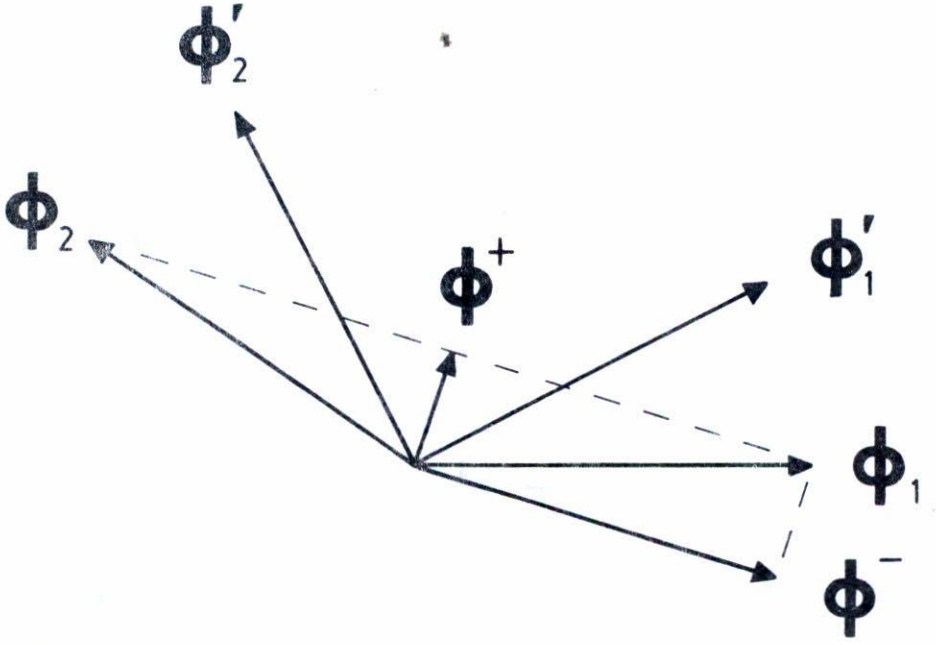

Fig. 7. Forming an appropriate linear combination of ϕ^+ and ϕ^- in order to obtain a symmetrically orthonormalized pair ϕ_1', ϕ_2', requires giving more and more weight to ϕ^+ as ϕ_2 becomes antiparallel to ϕ_1.

where

$$\underset{\sim}{S} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{29}$$

and $\lambda^{\pm}$ are given by (17). It should be noticed that $-\underset{\sim}{S}$ is also an admissible solution, but we need not consider it separately. Choosing $\underset{\sim}{U} = \underset{\sim}{S}$ in Eq. (7),

$$\underset{\sim}{0} = g^{-\frac{1}{2}} \cdot \underset{\sim}{S} \quad , \tag{30}$$

from (5) we obtain

$$\underset{\sim}{\phi}'_1 = (\sqrt{2}\,\lambda^+)\,\underset{\sim}{\phi}^+ \quad , \qquad \underset{\sim}{\phi}'_2 = (-\sqrt{2}\,\lambda^-)\,\underset{\sim}{\phi}^- \quad , \tag{31}$$

which correspond to Lowdin's canonical orthonormalization[2]. Its outstanding peculiarity is that it gives equal weight to both primitive vectors in the construction of the new ones, that is

$$\begin{aligned} |\underset{\sim}{\phi}_1 \cdot \underset{\sim}{\phi}'_1| &= |\underset{\sim}{\phi}_2 \cdot \underset{\sim}{\phi}'_1| = |\lambda^+/\sqrt{2}| \quad , \\ |\underset{\sim}{\phi}_1 \cdot \underset{\sim}{\phi}'_2| &= |\underset{\sim}{\phi}_2 \cdot \underset{\sim}{\phi}'_2| = |\lambda^-\sqrt{2}| \quad . \end{aligned} \tag{32}$$

Due to its consequence for the localization this property should be compared with the one

$$\underset{\sim}{\phi}_1 \cdot \underset{\sim}{\phi}'_1 = \underset{\sim}{\phi}_2 \cdot \underset{\sim}{\phi}'_2 \quad , \qquad \underset{\sim}{\phi}_1 \cdot \underset{\sim}{\phi}'_2 = \underset{\sim}{\phi}_2 \cdot \underset{\sim}{\phi}'_1 \quad , \tag{33}$$

corresponding to the symmetric orthonormalization.

When ϕ_1 and ϕ_2 are atomic orbitals centered on different sites the property Eq. (33) implies that the symmetric orthonormalization is the one which incorporates into each new orbital ϕ'_j the maximum amount of ϕ_j that is compatible with the preservation of the point symmetries, the latter being a feature of great importance in molecular and crystalline calculations. One thus obtains the most localized symmetry-preserving orthonormal basis. On the other hand, the canonical orthonormalization, by giving equal weight to both primitive vectors in the construction of the new ones, is the most delocalized of all possible orthonormal basis. These features are illustrated in Figs. 8 and 9 for two hydrogen orbitals centered $2a_0$ apart, where a_0 is the Bohr radius and S = 0.586453. For the sake of completeness we give in Fig. 10 the overlap S, and the angles γ and α corresponding to different separations R of the aforementioned orbitals, where S has been calculated using the standard formulas[6].

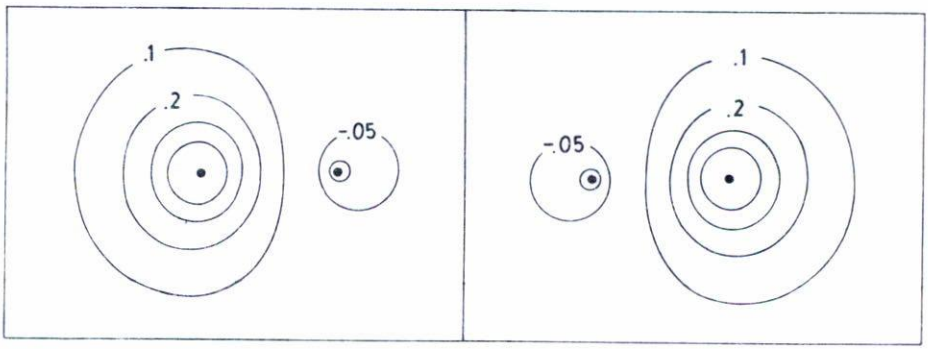

Fig. 8. Contours of the symmetrically orthonormalized orbitals corresponding to two hydrogen 1s orbitals whose nuclei (black dots) are $2a_0$ apart. The amplitude values are given in units of $a_0^{-3/2}$. The localization is the largest one compatible with the preservation of the reflection symmetry.

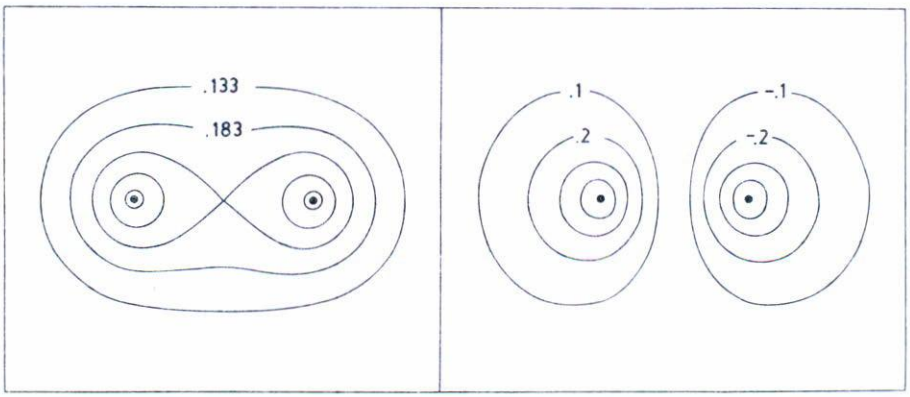

Fig. 9. Canonical orthonormalization of two hydrogen 1s orbitals $2a_0$ apart. All contours are equally spaced except the inner ones of the small lobes whose amplitudes are 0.1. Orbitals are as delocalized as is possible.

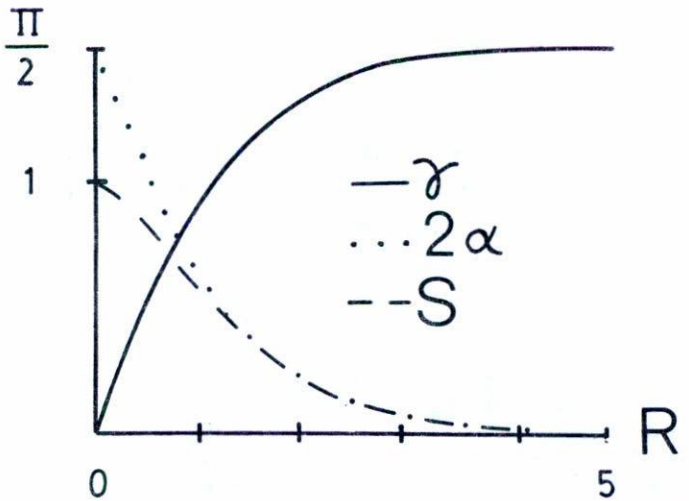

Fig. 10. Overlap S = cos γ, and angles γ and α = π/4 - γ/2 (symmetric orthonormalization) for two hydrogen 1s orbitals as a function of the distance R between centers. R is given in Angstroms and the angles in radians.

3. CONCLUSIONS

When restricting oneself to the plane, it becomes easy to give simple geometric interpretations of the properties of the Gram-Schmidt, symmetric and canonical orthonormalizations. Taking the primitive nonorthogonal vectors to be of unit length, the quantum-mechanical overlap S is then nothing but the cosine of the angle between vectors.

It is thus shown that the only method which preserves orthogonal relationships between vectors (as illustrated by a mirror line) is the symmetric one. This solution is also the one that minimizes the "distance" between the old and the new vectors, as given by Eq. (21). It is also clearly seen why some coefficients in the expansion of the latter in terms of the former blow up when approaching the limits S = ±1. This puts into sharper light the numerical problem of orthogonalizing quasi-linearly-dependent vectors.

The canonical orthonormalization turns out to be the one such that each new vector contains equal amounts of the old ones, as measured by the projection of the latter onto the former. This corresponds to the

maximum delocalization property of canonically orthonormalized orbitals, as opposed to the maximum localization of the symmetrically orthonormalized ones.

These easily grasped geometrical properties may be generalized to the n-dimensional case, thus providing a better understanding of the formerly abstract orthonormalization problem.

REFERENCES

1. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, New York, Interscience (1953) pp. 50, 3.
2. P.O. Löwdin, *Arkiv f. mat., astr.o.fysik*, 35A (1947) 9; P.O. Löwdin, *J. Chem. Phys.*, 18 (1950) 365; G.W. Pratt and S.F. Neustadter, *Phys. Rev.*, 101 (1956) 1248; P.O. Löwdin, *Adv. Phys.*, 5 (1956) 1; P.O. Löwdin, *Quantum Chemistry Group for Research in Atomic, Molecular and Solid State Theory*, Preliminary Research Report No. 213, Uppsala, (1968); P.O. Löwdin, *Adv. Quantum Chem.*, 5 (1970) 185; J.G. Aiken and H.B. Jonassen, *J. Chem. Phys.* 62 (1975) 2745 and 63 (1975) 2772.
3. J.C. Slater and G.F. Koster, *Phys. Rev.* 94 (1954) 1498.
4. G.A. Korn and T.M. Korn, *Mathematical Handbook for Scientists and Engineers*, 2nd edition, New York, McGraw-Hill (1961) p. 420.
5. B.C. Carlson and J.M. Keller, *Phys. Rev.* 105 (1957) 102.
6. R.S. Mulliken, C.A. Rieke, O. Orloff and H. Orloff, *J. Chem. Phys.* 17 (1949) 1248.