arXiv:quant-ph/9803073v1 26 Mar 1998

A new orthogonalization procedure with an extremal property

S. Chaturvedi*, A.K. Kapoor † and V. Srinivasan ‡

School of Physics

University of Hyderabad

Hyderabad - 500 046 (INDIA)

Abstract

Various methods of constructing an orthonomal set out of a given set of linearly independent vectors are discussed. Particular attention is paid to the Gram-Schmidt and the Schweinler-Wigner orthogonalization procedures. A new orthogonalization procedure which, like the Schweinler- Wigner procedure, is democratic and is endowed with an extremal property is suggested.

PACS No: 02.10.Sp, 31.15.+q

---

*e-mail:scsp@uohyd.ernet.in

†email:akksp@uohyd.ernet.in

‡e-mail:vssp@uohyd.ernet.in

Constructing an orthonormal set out of a given set of linearly independent Vectors is an age old problem. Among the many possible orthogonalization procedures, the two algorithmic procedures that have been extensively discussed and used in the literature are (a) the familiar Gram-Schmidt procedure [1] and (b) a procedure which is referred to as the Schweinler -Wigner procedure [2], particularly in the wavelet literature [3]. (This method is known among the chemists as the Löwdin orthogonalization procedure [4]. Mathematicians attribute it to Poincaré. Schweinler and Wigner themselves trace its origin to a work of Landshoff [5]. Eschewing the question of historically correct attribution, we shall continue to refer to it as the Schweinler-Wigner procedure) An intrinsic difference between the two procedures is that while the Gram-Schmidt procedure, by its very nature, requires one to select the linearly independent vectors sequentially, the Schweinler-Wigner procedures treats all the members of the set of linearly independent vectors democratically. The significance of the work of Schweinler and Wigner lies not in introducing a new orthogonalization method - the method was already known but rather in introducing a positive quantity m, to be defined shortly, which discriminates between various orthogonalization procedures. They showed that m is a maximum for the Schweinler-Wigner basis. In this letter we pose and answer the question as to what is the orthogonalization procedure which minimizes m. This new orthogonalization procedure, like the Schweinler-Wigner procedure, also turns out to be completely democratic in that it treats all the linearly independent vectors on the same footing. The quantity m was introduced by Schweinler and Wigner in a some what ad-hoc manner. We reformulate their procedure in a way so as that the quantity m appears in a natural way and can be useful in a wider context than that for which it was introduced. In particular, this reformulation enables us to quantify the notion of an orthonormal basis which brings any Hermitian operator in to a maximally off-diagonal form.

Let $v_1, \cdots, v_N$ denote a set of N linearly independent vectors. Let M denote the associated Gram matrix :$M_{ij} = (v_i, v_j)$. M is a positive definite Hermitian matrix. Define

$$\mathbf{z} = \mathbf{v}S \quad , \tag{1}$$

where S is an invertible matrix. Then

$$(z_i, z_j) = (S^\dagger M S)_{ij} \quad . \tag{2}$$

Requiring that $\mathbf{z}$ be an orthonormal basis amounts to requiring that

$$S^\dagger M S = I \quad i.e. \quad M^{-1} = S S^\dagger \quad . \tag{3}$$

Each such S defines an orthogonalization procedure. Two standard choices of S are

[1] Schweinler-Wigner Procedure

This procedure corresponds to the choice

$$S = U P^{-1/2} U^\dagger \quad , \tag{4}$$

where U is the matrix which brings M to a diagonal form P

$$U^\dagger M U = P \quad . \tag{5}$$

With this choice of S, which corresponds to taking the Hermitian square root of the matrix M, one has

$$\mathbf{z} = \mathbf{v} U P^{-1/2} U^\dagger \quad . \tag{6}$$

On defining $\mathbf{w} = \mathbf{z}U$, one obtains the Schweinler-Wigner basis

$$\mathbf{w} = \mathbf{v} U P^{-1/2} \quad . \tag{7}$$

[2] Gram-Schmidt Orthogonalization Procedure

In this procedure S is chosen to be an upper triangular matrix T satisfying

$$M^{-1} = T T^\dagger \tag{8}$$

and the Gram-Schmidt basis is given by

$$\mathbf{y} = \mathbf{v} T \tag{9}$$

The two orthonormal bases **w** and **y** discussed above are related to each other by the following unitary transformation

$$\mathbf{y} = \mathbf{w} V^{(1)} \quad , \tag{10}$$

where $V^{(1)}$ is given by

$$V^{(1)} = P^{1/2} U^{\dagger} T \quad . \tag{11}$$

Schweinler and Wigner introduced a quantity $m(\mathbf{z})$ as follows

$$m(\mathbf{z}) = \sum_k \left(\sum_l \mid (z_k, v_l) \mid^2 \right)^2 \quad , \tag{12}$$

where **z** is any orthonormal basis. They further showed that $m(\mathbf{z})$ attains its maximum value $Tr(M^2)$ for $\mathbf{z} = \mathbf{w}$.

$$m_{max} = Tr(M^2) = m(\mathbf{w}) \quad . \tag{13}$$

For any other basis **z**, related to **w** by a unitary transformation V

$$\mathbf{z} = \mathbf{w} V \quad , \tag{14}$$

the value of $m(\mathbf{z})$ is given by

$$m(\mathbf{z}) = \sum_k \left(\left(V^{\dagger} P V \right)_{kk} \right)^2 \quad . \tag{15}$$

In particular, for the Gram-Schmidt basis **y**, one finds that

$$m(\mathbf{y}) = \sum_k \left(\left(T^{\dagger} T \right)^{-1}_{kk} \right)^2 \quad . \tag{16}$$

A natural question to ask is as to what is the orthonormal basis which minimizes $m(\mathbf{z})$. On applying Cauchy inequality to (15), it follows that

$$\begin{aligned} m(\mathbf{z}) &= \sum_k \left(\left(V^{\dagger} P V \right)_{kk} \right)^2 \\ &\geq \frac{1}{N} \left(\sum_k (V^{\dagger} P V)_{kk} \right)^2 \\ &\geq \frac{1}{N} (Tr M)^2 \quad . \end{aligned} \tag{17}$$

Equality sign holds if and only if $(V^\dagger PV)_{kk} = c$ independent of k i.e

$$\sum_l P_l \mid V_{kl} \mid^2 = \frac{1}{N}(P_1 + \cdots + P_N) \tag{18}$$

This requires that

$$\mid V_{kl} \mid^2 = \frac{1}{N} \quad , \tag{19}$$

for all k and l. The matrix elements of the unitary matrix $V^{(2)}$ satisfying the above equation are thus given by

$$V^{(2)}_{kl} = \frac{1}{\sqrt{N}} \exp\left[\frac{2\pi i(k-1)(l-1)}{N}\right] \quad . \tag{20}$$

The matrix $V^{(2)}$ is thus just the character table of the cyclic group C_N. Thus the basis $\mathbf{x}$, for which $m(\mathbf{z})$ attains its minimum value $m_{min} = (1/N)(TrM)^2$, is related to the Schweinler-Wigner basis $\mathbf{w}$ as follows

$$x_l = \frac{1}{\sqrt{N}} \sum_k \exp\left[\frac{2\pi i(k-1)(l-1)}{N}\right] w_k \quad . \tag{21}$$

This basis also treats all the linearly independent vectors $\mathbf{v}$ democratically like the Schweinler-Wigner basis.

The quantity $m(\mathbf{z})$ appears to have been introduced by Schweinler and Wigner in a rather ad-hoc way. At least no particular motivation for introducing it appears in their work. We now reformulate their work in a way that this quantity appears naturally. Consider the Hermitian operator

$$\mathcal{M} = \sum_k v_k v_k^\dagger \quad . \tag{22}$$

In an arbitrary orthonormal basis $\mathbf{z}$ one can write

$$Tr(\mathcal{M}^2) = \sum_{\substack{lm \\ l\neq m}} \mid (z_l, \mathcal{M} z_m) \mid^2 + \sum_l \mid (z_l, \mathcal{M} z_l) \mid^2 \quad . \tag{23}$$

The second term on the rhs is easily seen to be the same as $m(\mathbf{z})$ in (12). From this perspective it is immediately obvious that the basis which maximizes $m(\mathbf{z})$ is the one in which $\mathcal{M}$ is diagonal. This is just the Schweinler-Wigner basis as can also be directly verified. Thus the Schweinler-Wigner basis is simply the eigenbasis of the operator $\mathcal{M}$ and if the eigenvalues of $\mathcal{M}$ are all distinct then this basis is essentially unique. Further, since $Tr(\mathcal{M}^2)$ is independent of the choice of basis, it is clear from (23) that the basis which minimizes $m(\mathbf{z})$ maximizes

$$n(\mathbf{z}) \equiv \sum_{\substack{lm \\ l \neq m}} \mid (z_l, \mathcal{M} z_m) \mid^2 \quad . \tag{24}$$

The quantity $n(\mathbf{z})$ therefore provides a quantitative measure of the off-diagonality of the operator $\mathcal{M}$ in the $\mathbf{z}$ basis. The new orthonormal basis proposed in this work is thus the one in which $\mathcal{M}$ is maximally off-diagonal.

Acknowledgements

We are grateful to Prof G.S. Agarwal, Prof Vipin Srivastava and to Dr. M. Durgaprasad for numerous discussions.

References

[1] See, for instance, Gantmacher F C 1960 *The Theory of Matrices* Vol 1 (New York : Chelsea)

[2] Schweinler H C and Wigner E P 1970 J. Math. Phys. **11** 1693

[3] See, for instance, Chui C K 1992 *Wavelet Analysis and its Applications* (San Diego CA : Academic Press)

[4] Löwdin P O 1950 J. Chem. Phys. **18** 365 ; 1956 Adv. Phys. **5** 172 ; 1992 Adv. Quant. Chem. **23**, 84

[5] Landshoff R 1936 Z. Physik **102**, 201