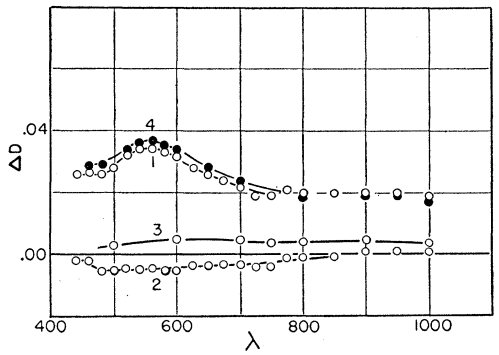

FIG. 4. Change in optical density (ΔD) *vs* wavelength (λ) in millimicrons for crystal exposed at -28°C. Curve 1—Exposed to 2×10^{16} quanta per square centimeter at 416 mμ (approximately 20% absorbed). Curve 2—Bleached by exposure to 1×10^{20} quanta per square centimeter at 580 mμ (less than 8% absorbed). Curve 3—After curve 2 the crystal was warmed to room temperature and then recooled to -28°C. Curve 4—Exposed to 1.5×10^{16} quanta per square centimeter at 416 mμ.

mental absorption band is a function of the temperature of the crystal.

Although efficient selective bleaching of the induced band does occur near the peak of the band, long illumination with 580-mμ light will ultimately bleach the crystal completely. This is most striking in the case of a crystal darkened at -28°C. Here a 60-mμ band of illumination completely bleached a band of color which extended beyond the 600-mμ range investigated for this paper.

It is also possible that several different kinds of absorption centers may form in the crystal if for no other reason than the "bleaching" light of wavelength 580 mμ can itself produce darkening at room temperature. There exists some evidence that at least some of the darkening produced by extended exposure to 580-mμ light is due to a surface colloidal deposit of silver.

It might be profitable to extend the investigation to include a wider range of temperatures and to try the darkening and bleaching with a series of different wavelengths where the radiation is more nearly monochromatic.

ACKNOWLEDGMENT

The author wishes to thank Dr. F. C. Brown for many discussions of this work and the Research Corporation for its vital financial support.

PHYSICAL REVIEW VOLUME 105, NUMBER 1 JANUARY 1, 1957

Orthogonalization Procedures and the Localization of Wannier Functions*

B. C. CARLSON AND JOSEPH M. KELLER
Institute for Atomic Research and Department of Physics, Iowa State College, Ames, Iowa
(Received September 26, 1956)

The method of "symmetric orthonormalization" is shown to have a least-squares property: it constructs those unique orthonormal functions which minimize the sum of squared distances (in Hilbert space) between each initial function and a corresponding function of the orthonormal set. The localization of Wannier functions is a consequence of this property, since they can be obtained from localized atomic orbitals by symmetric orthonormalization. The theorem further implies an optimal resemblance of Wannier functions to atomic orbitals.

IN the many-electron problems of molecular or solid-state physics, it is often desirable to work with orthogonal one-electron functions. But the natural starting functions may be, for instance, atomic orbitals that are localized about different centers and overlap one another to some extent. The familiar Schmidt procedure of orthogonalizing the functions one at a time in successive steps leads to orthonormal functions that may bear successively less resemblance to the starting functions and that depend, moreover, on the order in which the starting functions are selected. Alternatively, one can sometimes use the symmetry properties of the system to construct functions that are orthogonal by virtue of having different symmetries; functions of this kind (molecular orbitals or Bloch functions) are spread out over the whole molecule or lattice. A third procedure, which shares with the second the advantage of treating all the starting functions on an equal footing, is the "symmetric orthogonalization" originated by Landshoff and Löwdin.[1] The orthonormal functions obtained by this method, for example the Wannier functions, are known to resemble the initial atomic orbitals in being localized about the lattice points (provided that the atomic orbitals do not overlap too much).[2]

The purpose of this note is to observe that the third procedure has a simple geometrical meaning that helps

* This work was performed in the Ames Laboratory of the U. S. Atomic Energy Commission.

[1] P. Löwdin, Advances in Phys. 5, 1 (1956), discusses orthogonalization procedures extensively and gives references to earlier papers in which the symmetric procedure was partially or completely formulated.

[2] G. Wannier, Phys. Rev. 52, 191 (1937). Also see reference 1, Eq. (3.78).

one to understand why it leads to localized functions. Specifically, the symmetric procedure constructs those unique orthonormal functions which minimize the sum of the squares of the distances between each initial function f_i and a corresponding function φ_i of the orthonormal set. Distance means distance in Hilbert space; that is, the minimized sum is $\sum_i \int |\varphi_i - f_i|^2 d\tau$. The localization of the Wannier functions may be said to result from the fact that they approximate the atomic orbitals more closely in a least-squares sense than do the functions of any other orthonormal set.

The theorem to be proved is the following statement. Given a set of n linearly independent complex functions (or vectors) f_i $(i=1, 2, \cdots, n)$ with finite norms $\|f_i\| = (\int |f_i|^2 d\tau)^{\frac{1}{2}}$, then those orthonormal functions φ_i which minimize $\sum_i \|\varphi_i - f_i\|^2$ are determined uniquely to be $\varphi_i = \sum_j f_j (\Delta^{-\frac{1}{2}})_{ji}$, where Δ is the positive definite Hermitean matrix with elements $\Delta_{ij} = (f_i, f_j) = \int \bar{f}_i f_j d\tau$ and $\Delta^{-\frac{1}{2}}$ is the inverse of the (unique) positive definite Hermitean square root of Δ.

To prove the theorem, we choose any complete set of orthonormal functions φ_α $(\alpha=1, 2, \cdots)$ and expand $f_i = \sum_\alpha \varphi_\alpha F_{\alpha i}$. Summation over the Greek index α will always extend from one to infinity (or to the dimensionality of the space in the case of a finite-dimensional vector space); summation over i or j will extend from one to n. The matrix with elements $F_{\alpha i}$ is a rectangular matrix with n columns; it satisfies $F^*F = \Delta$, since $\Delta_{ij} = (f_i, f_j) = \sum_\alpha \bar{F}_{\alpha i} F_{\alpha j}$. The expansion of $\varphi_i - f_i = \sum_\alpha \varphi_\alpha \times (E_{\alpha i} - F_{\alpha i})$ involves a similar rectangular matrix E with elements $E_{\alpha i} = \delta_{\alpha i}$; clearly $E^*E = 1$ is the n by n unit matrix. The sum of squares is then

$$\sum_i \|\varphi_i - f_i\|^2 = \sum_{i\alpha} |F_{\alpha i} - E_{\alpha i}|^2 = \mathrm{Tr}[(F-E)^*(F-E)] = \mathrm{Tr}(\Delta + 1 - E^*F - F^*E).$$

We now ask how the first n functions of the orthonormal set should be chosen in order to minimize this sum of squares or, equivalently, what matrix F will maximize the trace of $E^*F + F^*E$ subject to the condition that $F^*F = \Delta$ is fixed.

Because the f_i are linearly independent, the Hermitean matrix Δ is positive definite[3] and has a unique[4] positive definite Hermitean square root, $\Delta^{\frac{1}{2}}$. Let R be any Hermitean square root of $\Delta^{\frac{1}{2}}$; positive definiteness is not required of R but it is required of $R^2 = \Delta^{\frac{1}{2}}$ in order that the (nonsingular) matrix R may be Hermitean. If we define a rectangular matrix $G = FR^{-1} - ER$, then it follows from $\Delta = R^4$ that

$$G^*G = R^{-1}\Delta R^{-1} + R^2 - RE^*FR^{-1} - R^{-1}F^*ER = 2\Delta^{\frac{1}{2}} - RE^*FR^{-1} - R^{-1}F^*ER.$$

The trace of G^*G is

$$\sum_{\alpha i} |G_{\alpha i}|^2 = 2\,\mathrm{Tr}\Delta^{\frac{1}{2}} - \mathrm{Tr}(E^*F + F^*E) = \sum_i \|\varphi_i - f_i\|^2 - \mathrm{Tr}[(\Delta^{\frac{1}{2}} - 1)^2].$$

If we choose $G=0$, or $F = ER^2 = E\Delta^{\frac{1}{2}}$, then $\sum_i \|\varphi_i - f_i\|^2$ is minimized without violating the condition $F^*F = \Delta$. The first n rows of this matrix F are the n by n matrix $\Delta^{\frac{1}{2}}$, and all other rows are identically zero. Thus $\varphi_1, \cdots, \varphi_n$ should be chosen to lie in the n-dimensional linear manifold spanned by the f_i, and they should be so chosen within this manifold that $f_i = \sum_j \varphi_j (\Delta^{\frac{1}{2}})_{ji}$. The solution of these equations is $\varphi_i = \sum_j f_j (\Delta^{-\frac{1}{2}})_{ji}$, which is precisely the recipe for symmetric orthogonalization.

It is interesting that the maximum value of the sum of squares is attained when the orthonormal functions again lie in the manifold spanned by the f_i and are just the negatives of the previous solutions. The proof is similar, with G replaced by $G' = FR^{-1} + ER$.

The least-squares property of the symmetric method suggests an extension to weighted least-squares orthogonalization: given a set of weights w_i, we may ask for orthonormal functions φ_i that minimize $\sum_i w_i \|\varphi_i - f_i\|^2$. More generally, if the matrix W is Hermitean and nonsingular (but not necessarily positive definite), then $\sum_{ij} W_{ji}(\varphi_i - f_i, \varphi_j - f_j)$ is minimized by choosing $\varphi_i = \sum_j f_j [W(W\Delta W)^{-\frac{1}{2}}]_{ji}$, where the square root is taken to be positive definite. A proof is easily constructed by defining $G = FW(W\Delta W)^{-\frac{1}{2}} - E(W\Delta W)^{\frac{1}{2}}$ and proceeding as before.

A further comment is that the f_i can always be orthogonalized by transformation with a unitary matrix. Since the Hermitean matrix Δ can be diagonalized by a unitary transformation, we can write $\Delta^{-\frac{1}{2}} = U\Lambda^{-\frac{1}{2}}U^{-1}$, where U is unitary (but not in general unique) and $\Lambda^{-\frac{1}{2}}$ is diagonal with positive diagonal elements. The functions $\psi_i = \sum_j f_j U_{ji}$ are then orthogonal, as asserted above, but not in general normalized. They become orthonormal if U is replaced by $U\Lambda^{-\frac{1}{2}}$; they remain orthonormal and acquire the least-squares property when the third factor U^{-1} is included. An example of these successive steps is the series of transformations, in the case of a crystal lattice, from atomic orbitals to unnormalized Bloch functions, then to normalized Bloch functions, and finally to Wannier functions.

We should like to thank Professor Klaus Ruedenberg for stimulating discussions.

Note added in proof.—Dr. Roland H. Good has pointed out to us that a paper by G. W. Pratt, Jr. and S. F. Neustadter, Phys. Rev. **101**, 1248 (1956), contains a result closely related to the one proved above. They show that if $|\Sigma_i(\varphi_i, f_i)|$ is stationary, then the φ's (except for the freedom of a common phase factor) are related to the f's by a Hermitean square root of Δ. The sum of squared distances used in the present paper involves the real part rather than the magnitude of $\Sigma_i(\varphi_i, f_i)$; maximizing the real part eliminates the freedom of a phase factor and also singles out the positive definite square root.

[3] $\sum_{ij} \bar{c}_i \Delta_{ij} c_j = \|\sum_i c_i f_i\|^2$ is positive unless all the coefficients c_i vanish.

[4] P. R. Halmos, *Finite Dimensional Vector Spaces* (Princeton University Press, Princeton, 1948), Sec. 66.