On the Factorization of Matrices.

On the Factorization of Matrices

by Norbert Wiener, South Tamworth (N. H.)

To Professor Plancherel, the founder of the precise theory of the Fourier integral and the inspirer of my work on harmonic analysis

§ 1. This note will deal primarily with binary matrices whose elements are functions of a variable ϑ which is to run between $(-\pi, \pi)$. It represents an extension of certain well-known theorems due to Szegö and the author, concerning scalar functions of ϑ. The fundamental theorem is the following:

Theorem 1. *Let $F(\vartheta)$ be non-negative and belong to Lebesgue class L over $(-\pi, \pi)$. Then a necessary and sufficient condition for us to be able to write*

$$F(\vartheta) = |\varphi(\vartheta)|^2 , \tag{1.01}$$

where

$$\varphi(\vartheta) = \sum_0^\infty a_n e^{in\vartheta} \tag{1.02}$$

and

$$\sum_0^\infty |a_n|^2 < \infty , \tag{1.03}$$

is that

$$\int_{-\pi}^{\pi} |\log F(\vartheta)|\, d\vartheta \tag{1.04}$$

be finite. It is then possible to choose the coefficients a_n in such a manner that

$$\Sigma\, a_n z^n \tag{1.05}$$

has no zeros inside the unit circle.

Let α be an arbitrary real number between 0 and 1. Let it be represented in the binary scale by the expression:

$$\alpha = .\, \alpha_1 \alpha_2 \alpha_3 \dots \tag{1.06}$$

Let these digits be re-numbered:

$$.\, \beta_0 \beta_1 \beta_{-1} \beta_2 \beta_{-2} \dots$$

and so on. Let

$$B_n(\alpha) = 2\beta_n - 1 . \tag{1.07}$$

It will follow that the transformation of α which changes $B_n(\alpha)$ into

$B_{n+1}(\alpha)$ for all values of α lying between 0 and 1, and all values of n, is a measure-preserving transformation T. We may write

$$B_{n+1}(\alpha) = B_n(T\alpha) \ . \tag{1.08}$$

This transformation T is not indeed well-defined for all values of α but is well-defined for all values of α with the exception of a set of measure 0.

If we start with any function $\varphi(\vartheta)$ belonging to L_2 and containing no negative frequencies, we can represent it, as I have said before, by the sequence of coefficients a_n where:

$$\sum_0^\infty |a_n|^2 < \infty \ . \tag{1.09}$$

Under these circumstances, it can be proved that

$$\sum_0^\infty a_n B_{-n}(\alpha) \tag{1.10}$$

will converge in the mean to a function of α which we shall call $f(\alpha)$. The function $f(\alpha)$ will then belong to L_2 over the interval $(0,1)$. If we consider the projection of any function $g(\alpha)$ belonging to L_2 on the closure of the set of

$$f(T^{-n}\alpha), f(T^{-n-1}\alpha), f(T^{-n-2}\alpha), \ldots \ , \tag{1.11}$$

this will converge in the mean to 0. It will obviously be the same as the projection of g on the closure of the set of functions $B_{-n}(\alpha)$, $B_{-n-1}(\alpha), \ldots$ That is, it will be the function

$$\sum_{\nu=0}^\infty B_{-\nu}(\alpha) \int_0^1 g(\beta) B_\nu(\beta)\, d\beta \ , \tag{1.12}$$

and will have as the integral of the square of its absolute value

$$\sum_{\nu=0}^\infty \left| \int_0^1 g(\beta)\, B_{-\nu}(\beta)\, d\beta \right|^2 \ . \tag{1.13}$$

This leads us immediately to the closely related

Theorem 2. *Let us assume in general that $f(\alpha)$ is any function whatever of the variable α which lies on $(0,1)$. Let T be any measure-preserving transformation of α into itself. Let the projection of $f(\alpha)$ on the set of functions*

$$f(T^{-n}\alpha), f(T^{-n-1}\alpha), \ldots \tag{1.14}$$

converge in the mean to 0 as n becomes infinite. Then there exists a function $h(\alpha)$ which is normalized which is linearly dependent on the set of functions

$$f(\alpha), \quad f(T^{-1}\alpha), \ldots$$

and which is orthogonal to all functions

$$f(T^{-1}\alpha), f(T^{-2}\alpha), \ldots . \tag{1.15}$$

It will follow that the functions $h(T^n\alpha)$ are a normal and orthogonal set, and it can be proved that $f(\alpha)$ will be equal to

$$f(\alpha) = \sum_0^\infty h(T^{-n}\alpha) \int_0^1 f(\beta)\overline{h(T^n\beta)}\, d\beta \tag{1.16}$$

as a limit in the mean. The function

$$\sum_0^\infty z^n \int_0^1 f(\beta)\overline{h(T^{-n}\beta)}\, d\beta \tag{1.17}$$

will be analytic inside the unit circle and will have no zeros there. Taken around any circle concentric with the unit circle but of smaller radius, the integral of the absolute square of this function will be uniformly bounded.

The statement in the hypothesis that $f(\alpha)$ is asymptotically orthogonal to the closure of

$$f(T^{-n}\alpha), f(T^{-n-1}\alpha), \ldots$$

as n becomes infinite is obviously a statement which merely concerns the autocorrelation coefficients

$$\int_0^1 f(T^n\alpha)\, \overline{f(\alpha)}\, d\alpha . \tag{1.18}$$

If then, these are of the form

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(\vartheta)\, e^{-in\vartheta}\, d\vartheta , \tag{1.19}$$

we can reduce this case to the particular case in which we have derived $f(\alpha)$ from $\varphi(\vartheta)$ by means of the $B's$.

§ 2. Now, let us start with two functions of class L_2, $f_1(\alpha)$, $f_2(\alpha)$. Parenthetically, let me remark that these both are to belong to L_2 and that we have one single transformation T of α into itself which preserves measure. Let the remote pasts of both f_1 and f_2 be asymptotically orthogonal to f_1 and f_2 which will be the case if $F_1(\vartheta)$ and $F_2(\vartheta)$ are respectively the functions belonging to L_2 with Fourier coefficients

$$\int_0^1 f_1(T^n\alpha) f_1(\alpha)\, d\alpha \tag{2.01}$$

and

$$\int_0^1 f_2(T^n\alpha) f_2(\alpha)\, d\alpha , \tag{2.02}$$

and let

$$\int_{-\pi}^{\pi} | \log F_1(\vartheta) | \, d\vartheta < \infty \ , \qquad \int_{-\pi}^{\pi} | \log F_2(\vartheta) | \, d\vartheta < \infty \ . \tag{2.03}$$

Under these circumstances we shall have two normalized functions $h_1(\alpha)$ and $h_2(\alpha)$ such that h_1 is linearly dependent on f_1 and $f_1(T^{-n}\alpha)$ and orthogonal to all functions $f_1(T^{-n}\alpha)$ where n is positive, and where h_2 will bear the same relation to $f_2(\alpha)$. We shall then have two normal and orthogonal set of functions $f_1(T^n\alpha)$ and $f_2(T^n\alpha)$, but there will not necessarily be any relation of orthogonality between these two sets.

Let us notice that if we put $F_{ij}(\vartheta)$ for the functions with Fourier coefficients

$$\int_0^1 f_i(T^n\alpha) f_j(\alpha) \, d\alpha \tag{2.04}$$

then

$$F_1(\vartheta) = F_{11}(\vartheta) \ , \tag{2.05}$$

and

$$F_2(\vartheta) = F_{22}(\vartheta) \ . \tag{2.06}$$

It is easy to prove that $F_{11}(\vartheta)$ and $F_{22}(\vartheta)$ are real and non-negative, while

$$F_{12}(\vartheta) = \bar{F}_{21}(\vartheta) \ . \tag{2.07}$$

Moreover,

$$\begin{vmatrix} F_{11}(\vartheta) & F_{21}(\vartheta) \\ F_{12}(\vartheta) & F_{22}(\vartheta) \end{vmatrix} \tag{2.08}$$

can be shown to be non-negative. Let us make the hypothesis

$$\int_{-\pi}^{\pi} \left| \log \begin{vmatrix} F_{11}(\vartheta) & F_{21}(\vartheta) \\ F_{12}(\vartheta) & F_{22}(\vartheta) \end{vmatrix} \right| d\vartheta < \infty \ . \tag{2.09}$$

Since we have made the supposition that the functions f_1 and f_2 belong to the class L_2, it is not difficult to prove that the functions $F_{ij}(\vartheta)$ all belong to the class L, so that the effective part of our assumption is

$$\int_{-\pi}^{\pi} \left| \log^- \begin{vmatrix} F_{11}(\vartheta) & F_{21}(\vartheta) \\ F_{12}(\vartheta) & F_{22}(\vartheta) \end{vmatrix} \right| d\vartheta < \infty \ . \tag{2.10}$$

Since however

$$\begin{vmatrix} F_{11}(\vartheta) & F_{21}(\vartheta) \\ F_{12}(\vartheta) & F_{22}(\vartheta) \end{vmatrix} = F_{11}F_{22} - F_{12}F_{21} = F_{11}F_{22} - | F_{12} |^2 \ , \tag{2.11}$$

it will follow that

$$\int_{-\pi}^{\pi} | \log^- F_{11}(\vartheta) F_{22}(\vartheta) | \, d\vartheta < \infty \tag{2.12}$$

from which we may conclude that

$$\left.\begin{aligned}\int_{-\pi}^{\pi} |\log F_{11}(\vartheta)|\, d\vartheta < \infty \\ \int_{-\pi}^{\pi} |\log F_{22}(\vartheta)|\, d\vartheta < \infty\end{aligned}\right\} \tag{2.13}$$

which are the assumptions we have previously made separately for

$$F_{11}(\vartheta) \quad \text{and} \quad F_{22}(\vartheta) \ .$$

§ **3.** I now wish to introduce a lemma of very general character concerning Hilbert space. It is the following:

Let H_1 be a closed subspace of Hilbert space and let H_2 be another such closed subspace. Then their common part $H_1 H_2$ will be a closed subspace of Hilbert space. If f is any vector in Hilbert space, and if $P_1 f$ is the projection of f on H_1 while $P_2 f$ is the projection of f on H_2, then the result of consecutive projection

$$P_1 f, P_2 P_1 f, \quad P_1 P_2 P_1 f, \ldots$$

will converge in the mean to the projection of f on $H_1 H_2$.

Let us note this H_1 contains two orthogonal spaces, one of which is $H_1 H_2$ while the other contains those functions in H_1 which are orthogonal to all functions in $H_1 H_2$. This other part we shall call H_1^*. Similarly, interchanging the rôles of H_1 and H_2, we separate every function of H_2 into a part lying in H_1 and a part orthogonal to all functions in $H_1 H_2$ which we call H_2^*. Then the successive projection of a vector on H_1 and H_2 will be given by its projection on $H_1 H_2$ plus the result of its successive projection on H_1^* and H_2^*. H_1^* and H_2^* will not necessarily be orthogonal to one another, but they will at any rate contain no vector other than 0 belonging to both. If therefore I can prove that when I have two closed subspaces of Hilbert space H_1^* and H_2^* not containing any vector in common except 0, then the result of consecutive projection of these two will converge in the mean to 0, I shall have established my lemma.

Now let $\varphi_n(\chi)$ be a set of normal and orthogonal functions belonging to H_1^* and closed on H_1^* and let $\psi_n(\chi)$ be a set of normal and orthogonal functions belonging to H_2^* and closed on H_2^*. Then if I start with any function $f(\chi)$ on H_1^* I can write it

$$\Sigma A_n \varphi_n(\chi) \ . \tag{3.01}$$

If I project this function on H_2^*, the projection will be

$$\sum_m \sum_n A_n \left[\int \varphi_n \overline{\psi_m}\right] \psi_m(x) \ ; \tag{3.02}$$

projecting this back on H_1^* I obtain

$$\sum_m \sum_n \sum_p A_n \left[\int \varphi_n \overline{\psi}_m \int \psi_m \overline{\varphi}_p\right] \varphi_p(x) \; . \tag{3.03}$$

The result of these repeated projections will be to change each function φ_n to

$$\Sigma\, Q_{pn} \varphi_p \tag{3.04}$$

where Q_{pn} will be

$$\sum_m \int \varphi_n \overline{\psi}_m \int \psi_m \overline{\varphi}_p \; . \tag{3.04}$$

That is Q_{pn} will satisfy the condition that

$$Q_{np} = \overline{Q}_{pn} \; . \tag{3.05}$$

The operator of double projection will have Hermitian coefficients and will be what is known as a self-conjugate operator. It will also be an operator which reduces the length of any known non-zero vector in H_1^*.

Well-known theorems of Hermann Weyl prove that such an operator will have a spectrum continuous or discrete. To transform any function in H_1 by such an operator, we expand it in the spectral functions, and change each function by a factor which is less than one in absolute value. It is easy to prove that such an operator, when repeated sufficiently often, will turn any vector of finite length into a vector of length as small as we choose.

Let us apply this lemma to the two spaces H_1 und H_2 consisting respectively of all functions of L_2 orthogonal to the functions $h_1(T^{-n}\alpha)$ and $h_2(T^{-n}\alpha)$. To form the projections of $h_1(\alpha)$ and $h_2(\alpha)$ on this space is essentially the same thing as taking the projections of f_1 and f_2 respectively on spaces which are respectively dependent on f_1 and its past, but orthogonal to its past and dependent on f_2 and its past and orthogonal to that past. Let me start with h_1 and find an expression for the part of h_1 which is orthogonal to the past of f_1 and f_2 and form the part of h_2 which is orthogonal to the past of f_1 and f_2. These functions we shall call respectively $k_1(\alpha)$, $k_2(\alpha)$.

We shall have for the projection of h_1 orthogonal to its own past h_1 itself, and $h_1(\alpha)$ will be our first approximation in the mean to $k_1(\alpha)$. We shall now take the part of h_1 which will be orthogonal to the past of h_2. This will be

$$h_1(\alpha) - \sum_{m=1}^{\infty} h_2(T^{-m}\alpha) \int_0^1 h_1(\beta)\overline{h_2(T^{-m}\beta)}\, d\beta \; . \tag{3.06}$$

We project again to find the part orthogonal to the part of H_1 where h_1 is orthogonal to its past and will need no new term so that only the second term must be taken care of. It is clear that the extra added term to make the third approximation will be given by

$$+\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} h_1(T^{-n}\alpha)\int_0^1 h_1(\beta)\overline{h_2(T^{-m}\beta)}\,d\beta\int_0^1 h_2(T^{-m}\beta)\overline{h_1(T^{-n}\beta)}\,d\beta\ . \tag{3.07}$$

The rule of continuing this series is now clear, and the terms will alternately contain $h_1(\alpha)$, the past of $h_2(\alpha)$, the past of $h_1(\alpha)$, and so on. The signs of the terms will alternate. The coefficient of the first term will contain one integral and one sign of summation, that of the second two integrals and two signs of summation, and so on. This series

$$\begin{aligned}
&h_1(\alpha)-\sum_{m=1}^{\infty} h_2(T^{-m}\alpha)\int_0^1 h_1(\beta)\overline{h_2(T^{-m}\beta)}\,d\beta\\
&+\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} h_1(T^{-m}\alpha)\int_0^1 h_2(T^{-m}\beta)\,\overline{h_1(T^{-m}\beta)}\,d\beta\int_0^1 h_1(\beta)\,\overline{h_2(T^{-n}\beta)}\,d\beta\\
&-\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{p=1}^{\infty} h_2(T^{-m}\alpha)\int_0^1 h_1(T^{-n}\beta)\,\overline{h_2(T^{-m}\beta)}\,d\beta\int_0^1 h_2(T^{-p}\beta)\,\overline{h_1(T^{-n}\beta)}\,d\beta\\
&\qquad\qquad\times\int_0^1 h_1(\beta)h_2(T^{-p}\beta)\,d\beta+\ldots
\end{aligned}$$

will be $k_1(\alpha)$. $k_1(\alpha)$ is then the part of $h_1(\alpha)$ which is orthogonal to the pasts of h_1 and h_2 so that

$$\begin{aligned}
&\int_0^1 |\,k_1(\alpha)\,|^2 d\alpha\\
&=\int_0^1 k_1(\alpha)\overline{k_1(\alpha)}\,d\alpha\\
&=\int_0^1 |\,h_1(\alpha)\,|^2 d\alpha-\sum_{m=1}^{\infty}\Big|\int_0^1 h_1(\alpha)\overline{h_2(T^{-m}(\alpha))}\,d\alpha\,\Big|^2\\
&-\sum_{m=1}^{\infty}\Big|\sum_{n=1}^{\infty}\int_0^1 h_1(\alpha)\overline{h_2(T^{-n}\alpha)}\,d\alpha\\
&\times\int_0^1 h_2(T^{-n}\alpha)\overline{h_1(T^{-m}\alpha)}\,d\alpha\,\Big|^2\ldots
\end{aligned} \tag{3.08}$$

Clearly

$$\int_0^1 |\,k_1(\alpha)\,|^2 d\alpha \tag{3.09}$$

is positive, and equally clearly

$$\int_0^1 |\,h_1(\alpha)\,|^2 d\alpha = 1\ . \tag{3.10}$$

Therefore

$$\int_0^1 | k_1(\alpha) |^2 d\alpha \tag{3.11}$$

lies between 0 and 1, and similarly

$$0 \leqslant \int_0^1 | k_2(\alpha) |^2 \, d\alpha \leqslant 1 \tag{3.12}$$

$k_1(\alpha)$ is that part of $h_1(\alpha)$ which is orthogonal to the pasts of both f_1 and f_2 while $k_2(\alpha)$ is that part of $h_2(\alpha)$ orthogonal to both pasts. Let us notice that

$$\int_0^1 k_i(T^{-n}(\alpha)) \overline{k_j(\alpha)} d\alpha \tag{3.13}$$

is always 0 if n is positive. From that and the measure-preserving character of T it results that

$$\int_0^1 k_i(T^n\alpha) \overline{k_j(T^m\alpha)} d\alpha \tag{3.14}$$

is 0 unless m and n are the same. As yet however, we know nothing in the case where m and n are the same, except that we may reduce this case to the case when both m and n may be given the value 0.

There are two cases which now present themselves. Either k_1 and k_2 have a relation of linear dependence or they do not. If they are linearly independent, they cannot, either of them, be equivalent to 0. Let us suppose that k_1 is not equivalent to 0. Then we can normalize it to obtain $q_1(\alpha)$. We then form

$$k_2(\alpha) - q_1(\alpha) \int_0^1 k_2(\beta) q_1(\beta) d\beta \; . \tag{3.15}$$

This function is obviously orthogonal to q_1. If it is equivalent to 0, k_1 and k_2 are not linearly independent. If it is not equivalent to 0, it can be normalized, and thus we obtain $q_2(\alpha)$. Then the functions $q_1(\alpha)$ and $q_2(\alpha)$ are such that $q_i(T^m\alpha)$ form a normal and orthogonal set, any two of them being orthogonal, unless both i and m agree.

Continuing on the assumption that k_1 and k_2 are linearly independent, we can express f_1 and f_2 in terms of this normal and orthogonal set. In proving this, we can establish that the formal series for $f_i(\alpha)$ is

$$\sum_{j=1,2} \sum_{n=1}^{\infty} q_j(T^{-n}\alpha) \int_0^1 f_i(\beta) \overline{q_j(T^{-n}\beta)} d\beta \; . \tag{3.16}$$

By studying the partial sums of this series and the difference between these partial sums and $f_i(\alpha)$, we can see that either the series converges in the mean to $f_i(\alpha)$, or we shall have the projection of $f_i(\alpha)$ on the remote past of f_1 and f_2 together not going to 0. Since the latter has been excluded, we shall have

$$f_i(\alpha) = \sum_{j=1,2} \sum_{n=1}^{\infty} q_j(T^{-n}\alpha) \int_0^1 f_i(\beta)\overline{q_j(T^{-n}\beta)}d\beta \ . \tag{3.17}$$

Under these conditions

$$\int_0^1 f_i(T^m)\overline{f_j(\alpha)}d\alpha$$

$$= \sum_{K=1,2} \sum_{n=0}^{\infty} \int_0^1 f_i(T^{m+n}\beta)\overline{q_K(\beta)}d\beta \int_0^1 \overline{f_j(T^n\beta)}\ q_K(\beta)d\beta \ . \tag{3.18}$$

Let us notice that the series

$$M_{ij}(\vartheta) = \sum_0^{\infty} e^{in\vartheta} \int_0^1 f_i(\beta)\overline{q_j(T^{-n}\beta)}d\beta \tag{3.19}$$

will converge in the mean to functions belonging to L_2, and that this will be equal to

$$\int_0^1 f_i(T^m\alpha)\overline{f_j(\alpha)d\alpha} = \frac{1}{2\pi}\int_{-\pi}^{\pi}[M_{i,1}(\vartheta)\ \overline{M_{j,1}(\vartheta)} + M_{i,2}(\vartheta)\overline{M_{j,2}(\vartheta)}]\,e^{im\vartheta}d\vartheta\,. \tag{3.20}$$

In other words, if we use matrix notation, the matrices whose Fourier coefficients are given by the autocorrelation matrices with elements belonging to L, can be factored into the matrix product

$$\underset{\sim}{M}\cdot\overline{\underset{\sim}{\tilde{M}}}\ , \tag{3.21}$$

where all the elements of $\underset{\sim}{M}$ are the boundary values on the unit circle of functions of class L_2 analytic inside the unit circle, and indeed where it will not be difficult to show that the determinants of these matrices have no 0's inside the unit circle.

The other case which we have not yet discussed is that in which there is a linear relation between k_1 and k_2. If there is such a linear relation, at least one of the functions g_1 or g_2 can be expressed linearly in terms of the other and the past of both. In other words, we have a relation such as $f_1(\alpha) = cf_2(\alpha) +$ a vector in the past of f_1 and f_2.

Under these circumstances

$$\int_0^1 f_1(T^n\alpha)\overline{f_i(\alpha)}d\alpha = c\int_0^1 f_2(T^n\alpha)\ \overline{f_i(\alpha)}d\alpha \tag{3.22}$$

plus something that may be approximated by a polynomial, always with the same coefficients, of the form

$$\int_0^1 f_2(T^{n-k}\alpha)\overline{f_j(\alpha)}\,d\alpha \tag{3.23}$$

and where the coefficients do not depend on n, but merely on i. It follows that if $\underset{\sim}{H}(\vartheta)$ is the Hermitian matrix, of which the autocorrelation coefficients are Fourier transforms, its elements will be such that

$$H_{1j}(\vartheta) = cH_{2j}(\vartheta) + \varphi_1(\vartheta)H_{1j}(\vartheta) + \varphi_2(\vartheta)H_{2j}(\vartheta)\ . \tag{3.24}$$

where φ_1 and φ_2 are free from singularity inside the unit circle. That is, the determinant

$$|\underset{\sim}{H}(\vartheta)| \tag{3.25}$$

will vanish identically inside of the unit circle, and therefore by a simple limit theorem, will vanish almost everywhere on the periphery. In other words, we have a situation which contradicts our assumption that

$$\int_{-\pi}^{\pi} |\log|\underset{\sim}{H}(\vartheta)||\,d\vartheta \tag{3.26}$$

is finite. We may sum up these results in the following words. *If the Hermitian matrix $\underset{\sim}{H}(\vartheta)$ has Fourier coefficients of the form*

$$\int_0^1 f_i(T^n\alpha)\overline{f_j(\alpha)}\,d\alpha \tag{3.27}$$

where f_1 and f_2 belong to L_2, and if

$$\int_{-\pi}^{\pi} |\log|\underset{\sim}{H}(\vartheta)||\,d\vartheta \tag{3.28}$$

converges, then we may write

$$\underset{\sim}{H}(\vartheta) = \underset{\sim}{M}(\vartheta)\,\underset{\sim}{\tilde{\bar{M}}}(\vartheta) \tag{3.29}$$

where the elements of $\underset{\sim}{M}$ belong to L_2 inside any smaller circle concentric with the unit circle, and converge in the mean to their value on the unit circle. Indeed the determinant of the matrix $\underset{\sim}{M}$ will be free from zeros inside the unit circle.

§ **4.** We wish now to establish two further things: one, that any Hermitian matrix of positive type for which the integral of the logarithm of the determinant converges, can be represented in the manner given above; and second, that if the integral of the logarithm diverges, the matrix cannot be factored in the manner indicated. In order to establish the first of these results, let us suppose that a Hermitian matrix $\underset{\sim}{H}$ can be written in the form

$$\underset{\sim}{H}(\vartheta) = \underset{\sim}{M}(\vartheta) \cdot \tilde{\bar{\underset{\sim}{M}}}(\vartheta) \tag{4.01}$$

where $\underset{\sim}{M}$ is a matrix belonging to L_2. This is what we shall mean by saying that H is Hermitian and of positive type. I now introduce a variable α which I represent as before in binary form, but I now split its digits into two sequences labelled from $(-\infty, \infty)$ according to the rule

$$\begin{aligned} \alpha &= \alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5 \ldots \\ &= \beta_0 \gamma_0 \beta_1 \gamma_1 \beta_{-1} \gamma_{-1} \beta_2 \gamma_2 \beta_{-2} \gamma_{-2} \cdots \end{aligned} \tag{4.02}$$

I write

$$B_n(\alpha) = 2\beta_n - 1 \; ; \qquad \Gamma_n(\alpha) = 2\gamma_n - 1 \; . \tag{4.02}$$

I introduce the transformation on α given by

$$T\alpha = .\beta_1 \gamma_1 \beta_2 \gamma_2 \beta_0 \gamma_0 \beta_3 \gamma_3 \beta_{-1} \gamma_{-1} \cdots \; . \tag{4.03}$$

I put

$$M_{ij,n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} M_{ij}(\vartheta)\, e^{-in\vartheta}\, d\vartheta \; . \tag{4.04}$$

I now define $f_i(\alpha)$ when i is one or two by

$$f_i(\alpha) = \sum_n \left(M_{i1,n} B_n(\alpha) + M_{i2,n} \Gamma_n(\alpha) \right) \; . \tag{4.05}$$

Then it will not be difficult to prove that $\underset{\sim}{H}(\vartheta)$ will have Fourier coefficients which can be written in the form

$$\int_0^1 f_i(T^n\alpha)\, \overline{f_j(\alpha)}\, d\alpha \; . \tag{4.06}$$

It remains to prove that if our logarithmic integral is infinite, no factorization can take place. However, if the factorization takes place and the said integral is infinite, then $\underset{\sim}{M}(\vartheta)$ will exist such that all the elements will belong to L_2 and will be boundary values of functions analytic inside the unit circle and

$$\int \left| \log \left| \underset{\sim}{M}(\vartheta) \right| \right| d\vartheta \tag{4.07}$$

are divergent. However, the determinant $|\underset{\sim}{M}(\vartheta)|$ will be a function of L_2 around the unit circle and without zeros inside the unit circle, and we need only to appeal to our scalar theorem to show the impossibility of the vector situation.

§ 5. Having established our factorization theorem for Hermitian matrices of positive type, let us examine some of the consequences of this for a more general type of matrix. Suppose that $\underset{\sim}{H}$ is Hermitian and of positive type, which simply amounts to assuming that $\underset{\sim}{H}$ can be written in the form

$$\underset{\sim}{H}(\vartheta) = \underset{\sim}{M}(\vartheta)\, \underset{\sim}{\tilde{\overline{M}}}(\vartheta) \ , \tag{5.01}$$

and that $\underset{\sim}{M}$ is an arbitrary matrix of class L_2. Let us notice that

$$| \underset{\sim}{H}(\vartheta) | = |(| \underset{\sim}{M}(\vartheta) |)|^2 \ , \tag{5.02}$$

so that

$$\int_{-\pi}^{\pi} | \log || \underset{\sim}{H}(\vartheta) || \,|\, d\vartheta < \infty \tag{5.03}$$

is equivalent to

$$\int_{-\pi}^{\pi} | \log | \underset{\sim}{M}(\vartheta) | \,|\, d\vartheta < \infty \ . \tag{5.04}$$

Then we may write that

$$\underset{\sim}{H}(\vartheta) = \underset{\sim}{M}^*(\vartheta) \cdot \underset{\sim}{\tilde{\overline{M}}}{}^*(\vartheta) \ , \tag{5.05}$$

where $\underset{\sim}{M}^*(\vartheta)$ is a function of L_2 around the unit circle which is the boundary value of a function free from singularities inside. Inside the unit circle it follows that

$$\underset{\sim}{M}^{-1}(\vartheta)\, \underset{\sim}{M}^*(\vartheta) \cdot \underset{\sim}{\tilde{\overline{M}}}{}^*(\vartheta)\, (\underset{\sim}{\tilde{\overline{M}}}{}^{-1}(\vartheta)) = \underset{\sim}{I} \ . \tag{5.06}$$

However, it is easy to prove that

$$(\underset{\sim}{\tilde{\overline{M}}}{}^{-1}) = (\underset{\sim}{\tilde{\overline{M}}})^{-1} \ . \tag{5.07}$$

Under these circumstances the matrix

$$\underset{\sim}{M}^{-1}(\vartheta)\, \underset{\sim}{M}^*(\vartheta) \tag{5.08}$$

will be a unitary matrix $\underset{\sim}{U}(\vartheta)$, such that

$$\underset{\sim}{U}(\vartheta) \cdot \underset{\sim}{\tilde{\overline{U}}}(\vartheta) = \underset{\sim}{I} \ . \tag{5.09}$$

It follows from this that

$$\underset{\sim}{M}(\vartheta) = \underset{\sim}{M}^*(\vartheta)\, \underset{\sim}{U}^{-1}(\vartheta) \ ; \tag{5.10}$$

or that any matrix $\underset{\sim}{H}$ with elements belonging to L_2, is the product of a matrix of the type $\underset{\sim}{M}^*(\vartheta)$ and a unitary matrix. If then we can prove that any unitary matrix can be factored into the product

$$\underset{\sim}{U}_1(\vartheta)\, \underset{\sim}{U}_2(\vartheta) \ , \tag{5.11}$$

where $\underset{\sim}{U}_1$ and $\underset{\sim}{U}_2$ are both unitary matrices, but where $\underset{\sim}{U}_1$ is the boundary value of a unitary matrix with no singularities inside the unit circle, and where $\underset{\sim}{U}_2$ is the boundary value of a unitary matrix with no singularities outside the unit circle, then we shall be able to prove that

$$\underset{\sim}{M}(\vartheta) = \underset{\sim}{M}^*(\alpha)\, \underset{\sim}{U}_2^{-1}(\vartheta)\, \underset{\sim}{U}_1^{-1}(\vartheta) \; . \tag{5.12}$$

Here the product of the first two factors is the boundary value of a function with no singularities inside the unit circle, and $\underset{\sim}{U}^{-1}(\vartheta)$ has no singularities outside the unit circle. Thus to establish a general factorization theorem for all matrices of type L, what remains is the discussion of factorization theorem for unitary matrices.

§ 6. Every unitary matrix can be written in the form of $e^{i\varepsilon \underset{\sim}{H}}$ and if such a matrix depends on ϑ, it can be written in the form $e^{i\varepsilon \underset{\sim}{H}(\vartheta)}$. There is no difficulty in showing that this can be done in such a way that the elements of $\underset{\sim}{H}(\vartheta)$ are bounded. Furthermore, we can write the matrix $\underset{\sim}{H}(\vartheta)$ in a Fourier series

$$\sum_{-\infty}^{\infty} e^{in\vartheta}\, \underset{\sim}{H}_n \; . \tag{6.01}$$

We shall put

$$\underset{\sim}{H}_1(\vartheta) = \sum_{1}^{\infty} \underset{\sim}{H}_n e^{in\vartheta} \; , \tag{6.02}$$

and

$$\underset{\sim}{H}_2(\vartheta) = \sum_{-\infty}^{0} \underset{\sim}{H}_n e^{in\vartheta} \; . \tag{6.03}$$

Then clearly all the elements of the matrices $\underset{\sim}{H}_1(\vartheta)$ and $\underset{\sim}{H}_2(\vartheta)$ will belong to the Lebesgue class L_2.

Now I am going to suppose that

$$e^{i\lambda\vartheta \underset{\sim}{H}(\vartheta)} = \underset{\sim}{U}_1(\lambda, \vartheta) \cdot \underset{\sim}{U}_2(\lambda, \vartheta) \; , \tag{6.04}$$

where $\underset{\sim}{U}_1$ and $\underset{\sim}{U}_2$ are the boundary values of unitary matrices respectively analytic inside the unit circle and outside the unit circle. Then

$$\begin{aligned} e^{i(\lambda + d\lambda)\underset{\sim}{H}(\vartheta)} &= \underset{\sim}{U}_1(\lambda, \vartheta)\,(\underset{\sim}{U}_2(\lambda, \vartheta))\,(1 + i\,d\lambda\, \underset{\sim}{\overline{\overline{H}}}(\vartheta)) \\ &= \underset{\sim}{U}_1(\lambda, \vartheta) \cdot \underset{\sim}{U}_2(\lambda, \vartheta) \cdot (1 + i d\lambda\, \underset{\sim}{H}(\vartheta)) \cdot (\overline{\underset{\sim}{U}}_2(\lambda, \vartheta)\, \underset{\sim}{U}_2(\lambda, \vartheta)) \; . \end{aligned} \tag{6.05}$$

Now let us put

$$\underset{\sim}{U}_2(\lambda, \vartheta) \cdot \underset{\sim}{H}(\vartheta) \cdot \overline{\underset{\sim}{U}}_2(\lambda, \vartheta) = \underset{\sim}{K}(\vartheta) \; . \tag{6.04}$$

Then $\underset{\sim}{K}(\vartheta)$ will be bounded and can be separated like $\underset{\sim}{H}$ into the sum:

$$\underset{\sim}{K}_1(\vartheta) + \underset{\sim}{K}_2(\vartheta) \ , \tag{6.05}$$

where $\underset{\sim}{K}_1$ and $\underset{\sim}{K}_2$ both consist of elements belonging to L_2 and where they are respectively boundary values of matrix functions inside and outside the unit circle. It then follows that

$$e^{i(\lambda+d\lambda)\underset{\sim}{H}(\vartheta)} = \underset{\sim}{U}(\lambda,\vartheta)(1+id\lambda\underset{\sim}{K}_1(\vartheta))\,(1+id\lambda\underset{\sim}{K}_2\vartheta)\cdot\underset{\sim}{U}_2(\lambda,\vartheta) \ . \tag{6.06}$$

That is

$$\left.\begin{aligned} \frac{d\underset{\sim}{U}_1(\lambda,\vartheta)}{d\lambda} &= i\,\underset{\sim}{U}_1(\lambda,\vartheta)\cdot\underset{\sim}{K}_1(\vartheta) \ ; \\ \frac{d\underset{\sim}{U}_2(\lambda,\vartheta)}{d\lambda} &= i\,\underset{\sim}{K}_2(\vartheta)\cdot\underset{\sim}{U}_2(\lambda,\vartheta) \ . \end{aligned}\right| \tag{6.07}$$

From this stage on the completion of the factorization theorem is easy. Not only are the elements of the $\underset{\sim}{K}$'s functions belonging to L_2, but they all belong uniformly to L_2. If we subdivide the range of λ from 0 to 1 into small parts, we can then easily obtain an estimate for the error in factorization which we get by assuming these small parts to be infinitesimal parts, and this error can be made as small as we wish by a sufficiently fine subdivision of the range $0 \leqslant \lambda \leqslant 1$. Thus, starting with the trivial factorization of the identity matrix, we arrive at the case where $\lambda = 1$, and we have factored

$$\underset{\sim}{U}(\lambda) = e^{i\varepsilon\underset{\sim}{H}(\vartheta)} \ . \tag{6.08}$$

Notice that in (5.09), we have factored our matrix with L_2 elements $\underset{\sim}{M}(\vartheta)$ in the form

$$\underset{\sim}{M}^*(\vartheta)\,\underset{\sim}{U}_2^{-1}(\vartheta)\,\underset{\sim}{U}_1^{-1}(\vartheta) \ ; \tag{6.09}$$

or what is the same thing, if c is any constant, depending on ϑ, we have factored

$$\underset{\sim}{M}(\vartheta)c(\vartheta) \quad \text{into} \quad \underset{\sim}{M}^*(\vartheta)\,\underset{\sim}{U}_2^{-1}(\vartheta)\,\underset{\sim}{U}_1^{-1}(\vartheta)c(\vartheta) \ . \tag{6.10}$$

From this it is easy to show that we have factored any matrix $\underset{\sim}{M}(\vartheta)$ with elements belonging to the Lebesgue class L_2 into the two matrices

$$\underset{\sim}{M}_1(\vartheta) = \underset{\sim}{M}^*(\vartheta)\,\underset{\sim}{U}_2^{-1}(\vartheta) \qquad \text{and} \qquad \underset{\sim}{M}_2(\vartheta) = \underset{\sim}{U}_1^{-1}(\vartheta)c \ ,$$

where $\underset{\sim}{M}_1$ and $\underset{\sim}{M}_2$ have their elements of the Lebesgue class L_2 and are boundary values of matrix functions respectively analytic inside and outside the unit circle. This solves the matrix factorization problem for the binary case. Our necessary and sufficient condition for the factorizability of $\underset{\sim}{M}(\vartheta)$ belonging to L will be

$$\int_{-\pi}^{\pi} | \log | | \underset{\sim}{M}(\vartheta) | | | d\vartheta < \infty \ . \tag{6.12}$$

The factorization problem for matrices of higher order follows exactly the same lines but involves a somewhat greater complication of detail. This complication is only conspicuous in the positive Hermitian case, where the Hilbert-space theorem on which we have rested our proof, must be called in several consecutive times.

Once the factorization theorem has been established, it is available for the discussion of the solution of systems of linear integral equations representing extensions of the Hopf-Wiener integral equations. The author intends to devote a further memoir to the discussion of equations of this type.

Such systems of equations have already been proved by several authors, including Professor Harold Freeman of the Massachusetts Institute of Technology to be of considerable value in the study of the mathematical problems of operational analysis, and particularly in the problem concerning the optimum distribution of tolerances in the construction of a machine or an operational system.

The author wishes to thank Professor Freeman for calling this fact to his attention. He also wishes to thank Dr. Masani of Bombay for showing him the complete scope of the factorization problem, and for pointing out that it is not confined to the positive Hermitian case. Nevertheless, the positive Hermitian case contains the center of the difficulty of the most general case.

(Received July 29, 1954.)