Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty — II

By H. J. LANDAU and H. O. POLLAK

(Manuscript received July 6, 1960)

The theory developed in the preceding paper[1] *is applied to a number of questions about timelimited and bandlimited signals. In particular, if a finite-energy signal is given, the possible proportions of its energy in a finite time interval and a finite frequency band are found, as well as the signals which do the best job of simultaneous time and frequency concentration.*

I. INTRODUCTION AND SUMMARY

It is a common experience in the communications field that one cannot simultaneously confine a function $f(t)$ and its Fourier transform $F(\omega)$ too severely. The most familiar statement of this phenomenon is the Heisenberg *uncertainty principle*: If we measure the time-spread T of $f(t)$ by

$$T^2 = \frac{\int_{-\infty}^{\infty} (t - t_0)^2 \,|f(t)|^2 \,dt}{\int_{-\infty}^{\infty} |f(t)|^2 \,dt}$$

and the frequency-spread Ω of $F(\omega)$ by

$$\Omega^2 = \frac{\int_{-\infty}^{\infty} (\omega - \omega_0)^2 \,|F(\omega)|^2 \,d\omega}{\int_{-\infty}^{\infty} |F(\omega)|^2 \,d\omega}$$

then, for any choice of t_0 and ω_0, $\Omega T \geqq \frac{1}{2}$. Thus T and Ω cannot, for any Fourier transform pair, be both small. Equality will hold if $f(t)$ [and hence $F(\omega)$] are gaussian, and t_0 and ω_0 are chosen as the *means* of $|f(t)|^2$ and $|F(\omega)|^2$ (in this case both zero). This result, while demonstrating that our experience with timelimiting and bandlimiting is indeed related to mathematical truth, does not succeed in providing a very good understanding of what is really happening. We should like to know just how close one can come to simultaneous limiting in both time and frequency, and what the price is that one has to pay. We need a sharper measure of the concentrations of $f(t)$ and $F(\omega)$ than that afforded by the above variances of $| f(t) |^2$ and $| F(\omega) |^2$, a measure which, if possible, will depend on the behavior of $f(t)$ in a given finite time interval, and of $F(\omega)$ in a given finite frequency band.

An early attempt to meet this need was made by L. A. MacColl, who around 1940 proved the following previously unpublished form of the uncertainty principle:

If

$$\frac{\int_{t_0}^{t_0+T} | f(t) |^2 \, dt}{\int_{-\infty}^{\infty} | f(t) |^2 \, dt} = \alpha_1$$

and

$$\frac{\int_{\omega_0}^{\omega_0+\Omega} | F(\omega) | \, d\omega}{\int_{-\infty}^{\infty} | F(\omega) | \, d\omega} = \alpha_2,$$

then

$$\Omega T > 2\pi\alpha_1\alpha_2^2. \tag{1}$$

This theorem does indeed emphasize the behavior of $f(t)$ and $F(\omega)$ in given finite intervals. The quantity α_1, representing the proportion of the total energy of $f(t)$ which is in the time-interval $(t_0, t_0 + T)$, is especially satisfying as a measure of the spread of $f(t)$; on the other hand, α_2 has no immediate physical interpretation. A further difficulty with (1) is that there are no functions for which equality can be achieved, although in practice the estimate is quite good.

A more useful form of the uncertainty principle would replace the above measure α_2 by the proportion of energy of $F(\omega)$ in a frequency band, that is, by a definition similar to that of α_1. This is done in the present paper. We shall see that if

$$\frac{\int_{t_0-T/2}^{t_0+T/2} |f(t)|^2\,dt}{\int_{-\infty}^{\infty} |f(t)|^2\,dt} = \alpha^2$$

and

$$\frac{\int_{-\Omega}^{\Omega} |F(\omega)|^2\,d\omega}{\int_{-\infty}^{\infty} |F(\omega)|^2\,d\omega} = \beta^2;$$

then

$$\Omega T \geqq \Phi(\alpha,\beta),$$

where $\Phi(\alpha,\beta)$ will be found explicitly, the inequality will be sharp and functions yielding equality will be given. The optimal functions $f(t)$ will always be real if, as in the above statement, the frequency band is centered at zero. The same inequality holds if the frequency band under study is not centered at zero, but then the optimal functions are, in general, complex-valued.

The simplest special case of our result arises if $\beta = 1$, so that *all* of $F(\omega)$ is contained in $|\omega| \leqq \Omega$, and $F(\omega) = 0$ for $|\omega| > \Omega$. The question "if α is given, what is the minimum ΩT?" can now be rephrased "if ΩT is given, what is the maximum α?" Let us introduce the following notation: The *square norm* of f is the total energy of f:

$$\| f \|^2 = \int_{-\infty}^{\infty} |f(t)|^2\,dt.$$

Timelimiting a function f produces a function Df which is f restricted to $|t| \leqq T/2$:

$$Df \equiv \begin{cases} f & \text{if } |t| \leqq T/2 \\ 0 & \text{if } |t| > T/2. \end{cases}$$

Bandlimiting a function f produces a function Bf whose Fourier transform agrees with the Fourier transform of f for $|\omega| \leqq \Omega$, and vanishes for $|\omega| > \Omega$:

$$Bf = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} F(\omega) e^{i\omega t}\,d\omega.$$

By writing

$$F(\omega) = \int_{-\infty}^{\infty} f(s)e^{-i\omega s}\, ds,$$

we see that an alternative expression for Bf is given by

$$Bf = \frac{1}{\pi}\int_{-\infty}^{\infty} f(s)\,\frac{\sin \Omega(t-s)}{t-s}\, ds.$$

It was shown in the preceding paper[1] that if a function is bandlimited and then timelimited its energy must be reduced by at least a factor λ_0, where λ_0 it the largest eigenvalue of the integral equation

$$\lambda f(t) = \frac{1}{\pi}\int_{-T/2}^{T/2} f(s)\,\frac{\sin \Omega(t-s)}{t-s}\, ds. \qquad (2)$$

If, in particular, a function is already bandlimited ($f = Bf$), then by this result $\| Df \|^2 \leqq \lambda_0$. This, now, is just the special case of the uncertainty principle which we have been seeking: If $\beta = 1$, then $\alpha \leqq \sqrt{\lambda_0}$.

In the sequel, we shall take a longer look at this formula and its significance; let us, however, state the full result for all values of α and β:

Theorem: There is a function f such that $\| f \| = 1$, $\| Df \| = \alpha$ and $\| Bf \| = \beta$, under the following conditions, and only under the following conditions:

1. If $\alpha = 0$, when $0 \leqq \beta < 1$.
2. If $0 < \alpha < \sqrt{\lambda_0}$, when $0 \leqq \beta \leqq 1$.
3. If $\sqrt{\lambda_0} \leqq \alpha < 1$, when $\cos^{-1}\alpha + \cos^{-1}\beta \geqq \cos^{-1}\sqrt{\lambda_0}$.
4. If $\alpha = 1$, when $0 < \beta \leqq \sqrt{\lambda_0}$.

The body of the present paper will cover the following sequence of topics: Section II will develop the properties of timelimited and bandlimited functions, and the geometric interpretation of these properties, which we require. Section III contains the proof of the quoted theorem, a discussion of the "best" functions, and a number of pertinent graphs and numerical examples. Section IV indicates possible extensions of the theory, and includes the interesting result that if a timelimited function d and a bandlimited function b are given, it is always possible to find a "smallest" function f so that $Df = d$ and $Bf = b$. Finally, Section V gives applications of the preceding theory to filter theory, data transmission and antenna theory.

II. SPACES OF TIMELIMITED FUNCTIONS AND BANDLIMITED FUNCTIONS

We are concerned, in the present paper, with the collection of functions $f(t)$ which are square-integrable on $(-\infty, \infty)$. These form a Hilbert space, denoted by $\mathfrak{L}^2$, in which the inner product (f,g) is defined by

$$(f,g) = \int_{-\infty}^{\infty} f(t)\overline{g(t)}\, dt,$$

and $\| f \|^2 = (f,f)$ as usual.

The collection of timelimited functions forms a linear subspace $\mathfrak{D}$ of $\mathfrak{L}^2$ so that if f_1 and f_2 are timelimited, so is $af_1 + bf_2$. Furthermore, $\mathfrak{D}$ is *complete*, which means that if we have a sequence of functions $\{f_n\}$, $f_n \in \mathfrak{D}$ and if $\| f_n - f_m \| \to 0$, then there is a function $f \in \mathfrak{D}$ such that $\| f - f_n \| \to 0$.

Exactly the same statements may be made about bandlimited functions; they form a complete linear subspace $\mathfrak{B}$ of $\mathfrak{L}^2$. The latter statement follows from the earlier one through the *Parseval relation* for Fourier transforms: If F and G are the Fourier transforms of f and g respectively, then

$$\int_{-\infty}^{\infty} f(t)\overline{g(t)}\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)\overline{G(\omega)}\, d\omega.$$

We shall call two functions f and g *orthogonal* if

$$(f,g) = 0.$$

Notice that Df and $f - Df$ are orthogonal, since each one vanishes where the other one does not; by the Parseval relation, Bf and $f - Bf$ are also orthogonal.

The inner product permits us to define the *angle* between two functions f and g as follows: By the Schwarz inequality, we know that

$$| (f,g) | \leqq \| f \| \cdot \| g \| ;$$

since

$$| \operatorname{Re}(f,g) | \leqq | (f,g) | ,$$

we know that

$$-1 \leqq \frac{\operatorname{Re}(f,g)}{\| f \| \cdot \| g \|} \leqq 1 .$$

We may thus define the *angle* $\theta(f,g)$ between the functions f and g by

$$\theta(f,g) = \cos^{-1} \frac{\operatorname{Re}(f,g)}{\|f\| \cdot \|g\|}.$$

The extreme values 0 and π for $\theta(f,g)$ can be reached only if f and g are proportional (so that equality holds in the Schwarz inequality) and (f,g) is real.

Suppose now that $f \in \mathfrak{B}$ and $g \in \mathfrak{D}$, and that neither function vanishes identically. What can we say about the angle between them? The angle can vanish only if for some constant k, $f = kg$. But since $\mathfrak{B}$ and $\mathfrak{D}$ are linear spaces, this would mean that f is *both* timelimited and bandlimited, and this is known to be impossible.† If, then, the angle cannot vanish, can it be arbitrarily small? This is the key question which shall occupy us for some time. Let us consider, first of all, a fixed function $f \in \mathfrak{B}$, and an arbitrary $g \in \mathfrak{D}$. We know that $\theta(f,g)$ cannot vanish; is $\theta(f,g)$ bounded away from zero? If there is a greatest lower bound for $\theta(f,g)$, is it assumed for some particular functions $g \in \mathfrak{D}$? In this case, the answers are quite simple, and are given by the following:

Lemma 1: If $f \in \mathfrak{B}$ is given, then

$$\inf_{g \in \mathfrak{D}} \theta(f,g) > 0.$$

This infimum equals

$$\cos^{-1} \frac{\| Df \|}{\| f \|},$$

and is assumed by $g = kDf$ for any positive constant k.

Proof: If g is any function in $\mathfrak{D}$, then

$$\operatorname{Re}(f,g) \leqq |(f,g)| = |(Df,g)|$$

since

$$f = f - Df + Df \qquad \text{and} \qquad (f - Df,g) = 0.$$

But

$$|(Df,g)| \leqq \| Df \| \cdot \| g \|,$$

† For then

$$f(t) = \int_{-\Omega}^{\Omega} F(\omega)\, e^{i\omega t}\, d\omega,$$

since $f \in \mathfrak{B}$, would be an analytic function of the complex variable t whose vanishing for $|t| > T$ would imply $f \equiv 0$.

so that

$$\frac{\mathrm{Re}(f,g)}{\|f\| \cdot \|g\|} \leqq \frac{\|Df\|}{\|f\|} = \frac{\mathrm{Re}(f,Df)}{\|f\| \cdot \|Df\|}.$$

Since $\cos\theta$ is monotone decreasing in $(0,\pi)$, it follows that

$$\theta(f,g) \geqq \theta(f,Df)$$

for any $g \in \mathfrak{D}$, with equality whenever g and Df are proportional. This proves the lemma.

We proceed now to the case of arbitrary $f \in \mathfrak{B}$ and $g \in \mathfrak{D}$. Let us say, for convenience, if

$$\inf_{\substack{f \in \mathfrak{B} \\ g \in \mathfrak{D}}} \theta(f,g)$$

is actually assumed by specific functions, that the spaces $\mathfrak{B}$ and $\mathfrak{D}$ form *a least angle*. We now have the following:

Theorem 1: There exists a least angle between $\mathfrak{B}$ and $\mathfrak{D}$. This angle equals $\cos^{-1}\sqrt{\lambda_0}$, and is assumed by $\psi_0 \in \mathfrak{B}$ and $D\psi_0 \in \mathfrak{D}$, where λ_0 is the largest eigenvalue of (2), and ψ_0 the corresponding eigenfunction.

Proof: By the preceding lemma,

$$\min_{g \in \mathfrak{D}} \theta(f,g) = \cos^{-1}\frac{\|Df\|}{\|f\|},$$

so that

$$\inf_{\substack{f \in \mathfrak{B} \\ g \in \mathfrak{D}}} \theta(f,g) = \inf_{f \in \mathfrak{B}} \cos^{-1}\frac{\|Df\|}{\|f\|} \tag{3}$$

and the infimum on the left of (3) will actually be assumed if the infimum on the right is. It was shown in the preceding paper[1] that any $f \in \mathfrak{B}$ may be expanded in a series, convergent in L^2 mean, of the eigenfunctions ψ_n of (2),

$$f = \sum_{n=0}^{\infty} a_n \psi_n .$$

Then

$$\|f\|^2 = \sum_{0}^{\infty} |a_n|^2;$$

since

$$Df = \sum_{n=0}^{\infty} a_n D\psi_n ,$$

it follows from the properties of $\{D\psi_n\}$ that

$$\| Df \|^2 = \sum | a_n |^2 \lambda_n .$$

Thus

$$\cos^{-1} \frac{\| Df \|}{\| f \|} = \cos^{-1} \left(\frac{\sum | a_n |^2 \lambda_n}{\sum | a_n |^2} \right)^{\frac{1}{2}}.$$

Since it was shown in the preceding paper[1] that $\lambda_n < \lambda_0$, if $n \geqq 1$, it follows that

$$\max \left(\frac{\sum | a_n |^2 \lambda_n}{\sum | a_n |^2} \right)$$

is achieved if $a_n = 0$ for $n \geqq 1$, so that the minimum possible value of

$$\cos^{-1} \frac{\| Df \|}{\| f \|},$$

namely $\cos^{-1} \sqrt{\lambda_0}$, is actually assumed if $f = \psi_0$, and $g = D\psi_0$. The theorem is proved.

We have thus found that the two subspaces $\mathfrak{B}$ and $\mathfrak{D}$ of $\mathfrak{L}^2$, which have no functions except 0 in common, actually have a minimum angle between them, so that, in fact, a timelimited function and a bandlimited function cannot even be very close together. With the aid of this result, as we shall see, the uncertainty principle which we are seeking will follow.

In preparation for the coming theorems, we must consider one further aspect of the spaces $\mathfrak{B}$ and $\mathfrak{D}$. How close do $\mathfrak{B}$ and $\mathfrak{D}$ together come to filling up all of $\mathfrak{L}^2$? The two specific questions which concern us are the following: (i) if $\{f_n\}$, $f_n = b_n + d_n$ is a Cauchy sequence† of functions in $\mathfrak{B} + \mathfrak{D}$, what can the limiting function f look like; and (ii) do there exist functions $f \in \mathfrak{L}^2$ orthogonal to both $\mathfrak{B}$ and $\mathfrak{D}$ (i.e., to every function in $\mathfrak{B}$ and $\mathfrak{D}$)? The answers to these questions are the subjects of the subsequent two lemmas.

Lemma 2: If $\{f_n\}$ is a Cauchy sequence of functions of the form $f_n =$

† A Cauchy sequence of functions is a sequence such that $\| f_n - f_m \| \to 0$, so that, by the completeness of Hilbert space, there exists a limiting function f such that $\| f - f_n \| \to 0$.

$d_n + b_n$ where $d_n \in \mathfrak{D}$ and $b_n \in \mathfrak{B}$ for each n, then the limiting function f is itself of the form $d + b$, where $d \in \mathfrak{D}$ and $b \in \mathfrak{B}$.

Proof: For each $f_n = d_n + b_n$, we may also write

$$f_n = (b_n - Db_n) + (Db_n + d_n).$$

Here $Db_n + d_n \in \mathfrak{D}$, while $b_n - Db_n \perp \mathfrak{D}$. It now follows from the fact that the f_n form a Cauchy sequence that the functions $b_n - Db_n$ do; for

$$\| f_n - f_m \|^2 =$$
$$\| b_n - Db_n - (b_m - Db_m) \|^2 + \| Db_n + d_n + Db_m + d_m \|^2,$$

so that

$$\| b_n - Db_n - (b_m - Db_m) \| \leqq \| f_n - f_m \| .$$

But now, since $\{b_n - Db_n\}$ forms a Cauchy sequence, so does $\{b_n\}$ itself. For

$$\| b_n - b_m \|^2 = \| D(b_n - b_m) \|^2 + \| (b_n - b_m) - D(b_n - b_m) \|^2,$$

and by Lemma 1,

$$\| D(b_n - b_m) \| \leqq \sqrt{\lambda_0} \, \| b_n - b_m \| ,$$

so that

$$\| b_n - b_m \|^2 \leqq \frac{\| b_n - Db_n - (b_m - Db_m) \|^2}{1 - \lambda_0}.$$

Since $\{b_n\}$ is now a Cauchy sequence, there is a function $b \in \mathfrak{B}$ such that

$$\| b - b_n \| \to 0.$$

Thus $\{f_n\}$ and $\{b_n\}$ both converge in norm, and hence so does $\{d_n\}$, and to a limiting function $d \in \mathfrak{D}$ for which

$$f = b + d.$$

We have thus shown that taking a limit of sums of functions from $\mathfrak{B}$ and $\mathfrak{D}$ gives us nothing new, but only, once again, a sum of functions in $\mathfrak{B}$ and $\mathfrak{D}$. We may abbreviate this by saying simply that $\mathfrak{B} + \mathfrak{D}$ is *closed*.

Lemma 3: There are infinitely many functions in $\mathfrak{L}^2$ which are orthogonal to $\mathfrak{B} + \mathfrak{D}$.

Proof: The functions

$$f_n = \begin{cases} 1 & \text{if } T/2 + n \leqq |t| \leqq T/2 + n + 1 \\ 0 & \text{elsewhere} \end{cases} \qquad n = 0, 1, 2, \cdots$$

are instances of functions *not* in $\mathcal{B} + \mathcal{D}$, since the portion of f_n in $|t| > T/2$ is not a piece of a bandlimited function. Lemma 2 permits us to write the best approximation to f_n from $\mathcal{B} + \mathcal{D}$ in the form $b_n + d_n$, where $b_n \in \mathcal{B}$ and $d_n \in \mathcal{D}$; then

$$f_n^* = f_n - b_n - d_n$$

are distinct functions in $\mathcal{L}^2$ which are orthogonal to $\mathcal{B} + \mathcal{D}$.

There are in fact, in some sense "many more" functions in $\mathcal{L}^2 - \mathcal{B} - \mathcal{D}$ than in $\mathcal{B} + \mathcal{D}$; we do not know, however, of any really convenient representation for such functions.

III. THE UNCERTAINTY PRINCIPLE

We begin by restating the theorem announced in Section I.

Theorem 2: There is a function $f \in \mathcal{L}^2$ such that $\| f \| = 1$, $\| Df \| = \alpha$ and $\| Bf \| = \beta$, under the following conditions, and only under the following conditions:

1. If $\alpha = 0$, when $0 \leqq \beta < 1$.
2. If $0 < \alpha < \sqrt{\lambda_0}$, when $0 \leqq \beta \leqq 1$.
3. If $\sqrt{\lambda_0} \leqq \alpha < 1$, when $\cos^{-1}\alpha + \cos^{-1}\beta \geqq \cos^{-1}\sqrt{\lambda_0}$.
4. If $\alpha = 1$, when $0 < \beta \leqq \sqrt{\lambda_0}$.

Proof: Let $\mathcal{G}$ be the family of functions $f \in \mathcal{L}^2$ with $\| f \| = 1$ and $\| Df \| = \alpha$, and let us, for each case of α, determine

$$\sup_{f \in \mathcal{G}} \beta = \sup_{f \in \mathcal{G}} \| Bf \| .$$

We shall also show, in each case, that any value of β less than the supremum can be realized by an appropriate function. Whether or not the supremum itself can be realized will vary from case to case.

Case 1. $\alpha = 0$. If $\alpha = 0$, the family $\mathcal{G}$ can contain no function with $\beta = 1$. For if $f \in \mathcal{G}$ with $\beta = 1$ we must have $f \in \mathcal{B}$, whence f is analytic and vanishes for $|t| < T/2$ only if $f \equiv 0$. This is a contradiction.

To show that $\mathcal{G}$ contains functions with values of β arbitrarily close to 1 we set

$$f^* = \frac{\psi_n - D\psi_n}{\sqrt{1 - \lambda_n}},$$

where λ_n and ψ_n are respectively an eigenvalue and corresponding eigenfunction of (2). We observe that $f^* \in \mathcal{G}$ and that $\beta = \| Bf^* \| = \sqrt{1 - \lambda_n}$. Since there exist eigenvalues λ_n arbitrarily small, there exist functions in $\mathcal{G}$ with values of β arbitrarily close to 1.

To find functions in $\mathcal{G}$ with values of β between those already covered, we consider $e^{i\rho t}f^*(t)$, which belongs to $\mathcal{G}$ since $\| e^{i\rho t}f^* \| = \| f^* \| = 1$ and $\| D\, e^{i\rho t}f^* \| = \| Df^* \| = \alpha$. For β we find

$$\beta = \| B\, e^{i\rho t} f^* \| = \left\{ \int_{-\rho-\Omega}^{-\rho+\Omega} | F^*(\omega) |^2 \, d\omega \right\}^{\frac{1}{2}},$$

where F^* is the Fourier transform of f^*. This quantity is continuous in ρ and approaches zero as $\rho \to \infty$, since $F^* \in \mathcal{L}^2$; thus $\mathcal{G}$ contains functions with all smaller values of β, except possibly $\beta = 0$.

A function f in $\mathcal{G}$ for which $\beta = 0$ must have the property that $Df = Bf = 0$; the existence of such functions was demonstrated in Lemma 3.

This completes the proof in Case 1; if we reverse B and D in the preceding arguments, we find that $\beta = 0$ is possible if and only if $0 \leqq \alpha < 1$; thus the minimum β in Cases 2 and 3 has also been established.

Case 2. $0 < \alpha < \sqrt{\lambda_0}$. Since $\lambda_n \to 0$ as $n \to \infty$, we can find an eigenvalue $\lambda_n < \alpha$. Let ψ_n be the corresponding eigenfunction, and consider

$$f^* = \frac{\sqrt{\alpha^2 - \lambda_n}\,\psi_0 + \sqrt{\lambda_0 - \alpha^2}\,\psi_n}{\sqrt{\lambda_0 - \lambda_n}}. \tag{5}$$

We have $f^* \in \mathcal{B}$, and $\| f^* \| = \| Bf^* \| = 1$, while a simple computation shows that $\| Df^* \| = \alpha$. This, then, covers the case $\beta = 1$; by picking $e^{i\rho t}f^*$, as in Case 1, we may obtain any $0 < \beta < 1$, and $\beta = 0$ is covered by the remark immediately preceding Case 2.

Cases 3 and 4. $\sqrt{\lambda_0} \leqq \alpha \leqq 1$. For a function $f \in \mathcal{G}$, let us find the closest point to f on the plane spanned by Df and Bf; we then can write

$$f = \lambda Df + \mu Bf + g, \tag{6}$$

with g orthogonal to both Df and Bf. Taking the inner product of (6) successively with f, Df, Bf and g, and using the fact that $f \in \mathcal{G}$, we obtain

$$\begin{aligned} 1 &= \lambda\alpha^2 + \mu\beta^2 + (g,f), \\ \alpha^2 &= \lambda\alpha^2 + \mu(Bf,Df), \\ \beta^2 &= \lambda(Df,Bf) + \mu\beta^2, \\ (f,g) &= (g,g). \end{aligned}$$

By eliminating (g,f), λ and μ from the above equations we find, for $\alpha\beta \neq 0$,

$$\beta^2 - 2\,\mathrm{Re}\,(Df,Bf) = -\alpha^2 + \left(1 - \frac{|(Df,Bf)|^2}{\alpha^2\beta^2}\right) - \|g\|^2\left(1 - \frac{|(Df,Bf)|^2}{\alpha^2\beta^2}\right). \tag{7}$$

We next set

$$\mathrm{Re}\,\frac{(Df,Bf)}{\|Df\|\cdot\|Bf\|} = \cos\theta.$$

The angle θ is that formed between $Df \in \mathfrak{D}$ and $Bf \in \mathfrak{B}$ so that, by Theorem 1,

$$\theta \geqq \cos^{-1}\sqrt{\lambda_0}\,. \tag{8}$$

Since

$$\alpha\beta\cos\theta = \mathrm{Re}(Df,Bf) \leqq |(Df,Bf)| \leqq \alpha\beta,$$

we have

$$0 \leqq 1 - \frac{|(Df,Bf)|^2}{\alpha^2\beta^2} \leqq 1 - \cos^2\theta. \tag{9}$$

Introducing θ into (7), completing the square on the left-hand side, and applying (9) we obtain

$$(\beta - \alpha\cos\theta)^2 \leqq (1 - \alpha^2)\sin^2\theta, \tag{10}$$

with equality if and only if $g = 0$ and (Df,Bf) is real. From (10) we find immediately

$$\beta \leqq \cos(\theta - \cos^{-1}\alpha),$$

whence by (8)

$$\beta \leqq \cos(\cos^{-1}\sqrt{\lambda_0} - \cos^{-1}\alpha), \tag{11}$$

or

$$\cos^{-1}\alpha + \cos^{-1}\beta \geqq \cos^{-1}\sqrt{\lambda_0}\,.$$

Equality in (11) is attained for the function

$$f^* = p\psi_0 + qD\psi_0\,, \tag{12}$$

with

$$p = \sqrt{\frac{1-\alpha^2}{1-\lambda_0}}$$

and (13)

$$q = \frac{\alpha}{\sqrt{\lambda_0}} - \sqrt{\frac{1 - \alpha^2}{1 - \lambda_0}},$$

since f^* satisfies all the conditions for equality in the above sequence of inequalities; the constants p and q are chosen so that $f \in \mathcal{G}$. As in Case 1, all smaller values of β, except possibly for $\beta = 0$, are attainable by the functions $e^{i\rho t}f^*(t)$ with suitable values of ρ, and, by the argument above, the family $\mathcal{G}$ contains functions with $\beta = 0$ as well, except when $\cos^{-1}\alpha = 0$. Thus, in Case 3, $\mathcal{G}$ is made up of functions for which β takes on all values for which

$$\cos^{-1}\alpha + \cos^{-1}\beta \geqq \cos^{-1}\sqrt{\lambda_0}\,.$$

If, however, $\alpha = 1$, we must exclude $\beta = 0$, so that we obtain in Case 4

$$0 < \beta < \sqrt{\lambda_0}\,.$$

The result of Theorem 2 is illustrated in Fig. 1, which shows the permissible region in the (α^2,β^2) plane for various values of $c = \Omega T/2$.

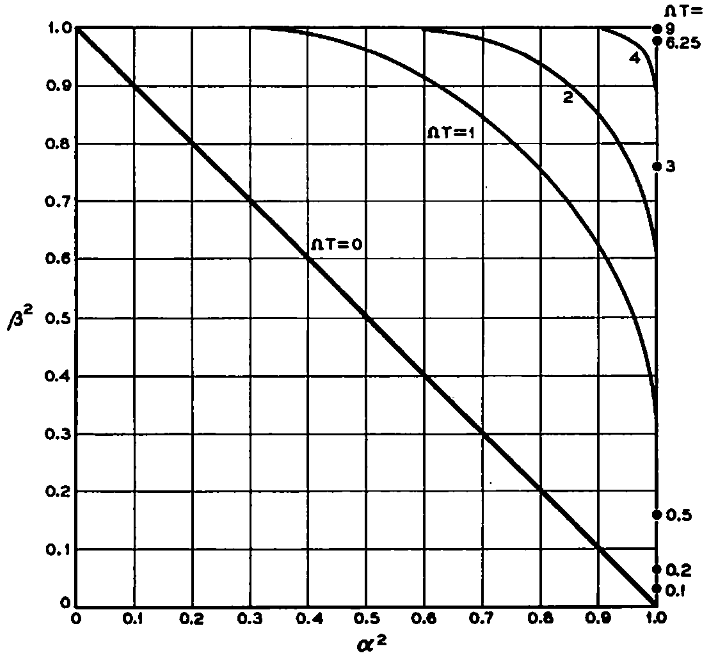

Fig. 1 — Possible combinations of α^2 and β^2 for different ΩT.

For each value of c, this region is bounded by the line segments

$$\begin{aligned} \alpha^2 &= 0 \quad \text{for} \quad 0 \leqq \beta^2 < 1, \\ \beta^2 &= 0 \quad \text{for} \quad 0 \leqq \alpha^2 < 1, \\ \alpha^2 &= 1 \quad \text{for} \quad 0 < \beta^2 \leqq \lambda_0(c), \\ \beta^2 &= 1 \quad \text{for} \quad 0 < \alpha^2 \leqq \lambda_0(c), \end{aligned}$$

and the curve $\cos^{-1}\alpha + \cos^{-1}\beta = \cos^{-1}\sqrt{\lambda_0(c)}$, which is labeled by the appropriate value of c.

An interesting phenomenon is brought up by the line $\alpha^2 + \beta^2 = 1$, which is labeled with $c = 0$. This labeling agrees with Theorem 2 in the following way:

If $\alpha^2 + \beta^2 \leqq 1$, then $\cos^{-1}\alpha + \cos^{-1}\beta \geqq \pi/2$, which automatically exceeds $\cos^{-1}\sqrt{\lambda_0}$ for any c, no matter how small. In physical terms, this observation states that if the proportions of energy of $f(t)$ in $|t| \leqq T/2$, and of $F(\omega)$ in $|\omega| \leqq \Omega$, add up to less than the total energy of $f(t)$, then we have really put no restraint on Ω and T, and an arbitrarily small ΩT product will still permit this distribution of energy. It is only when $\alpha^2 + \beta^2 > 1$, so that the energies in $|t| < T/2$ and in $|\omega| < \Omega$ add up to more than the total energy, that a nonzero lower bound on ΩT is implied.

Fig. 2 gives a detailed plot of what is essentially the top (or the right) edge of Fig. 1. We plot $\lambda_0(c)$, the maximum of α^2 if $\beta^2 = 1$, against c. We note that $\lambda_0(c) \to 1$ quite rapidly as $c \to \infty$; the approach is exponential, but the exact rate has not been proved. Fig. 2 also gives,

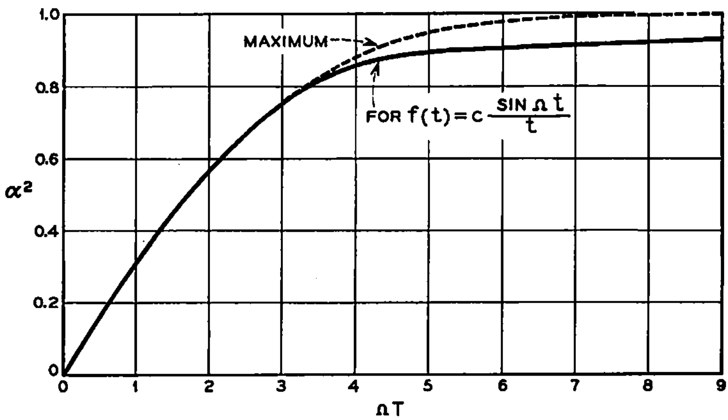

Fig. 2 — Possible α^2 if $\beta^2 = 1$.

for comparison, the proportion of energy in $| t | < T/2$ for the function

$$f(t) = \frac{\sin \Omega t}{t},$$

which has sometimes been "intuitively" considered as the bandlimited function which is as concentrated in time as possible. For small λ_0, it appears, $f(t)$ is indeed essentially as good as the optimal function; if, however, we wish to achieve a proportion of energy like 92 per cent, we see that $\Omega T = 4.5$ suffices, while use of $(\sin \Omega t)/t$ would require $\Omega T = 8.5$. For a proportion of 99 per cent, the minimal ΩT is 6.25, while $(\sin \Omega t)/t$ would require a value of ΩT of about 30.

Let us consider one more numerical example. If values of $\alpha^2 = 0.977$ and $\beta^2 = 0.96$ are desired, what are the minimum ΩT, and the corresponding optimal function? From $\cos^{-1}\alpha + \cos^{-1}\beta = \cos^{-1}\sqrt{\lambda_0}$ we find $\lambda_0 = 0.88$, so that $\Omega T = 4$, or $c = 2$. If, now, $\psi_0(t)$ is the first eigenfunction corresponding to $c = 2$, then, by (12) and (13), the optimal function (see Fig. 3) is $0.578\psi_0 + 0.465D\psi_0$. It is thus not a continuous function of t but has jumps at $t = \pm T/2$; this is characteristic of all of our optimization problems except for the special case $\beta^2 = 1$.

A note on previous work in the direction of Theorem 2. The connection between the extremum problem for $\beta^2 = 1$ and the largest eigenvalue of (2) was noted by Chalk[2] and Gurevich,[3] both of whom found the appearance of the optimal function without analytic solution

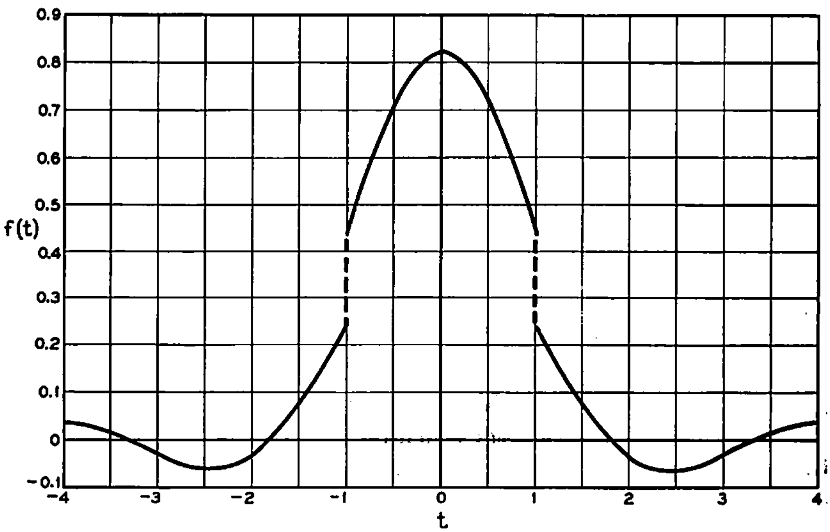

Fig. 3 — Plot of optimal $f(t)$ for $\alpha^2 = 0.977$, $\beta^2 = 0.96$, $T/2 = 1$.

of the integral equation; the latter also plotted the largest eigenvalue. The set of eigenfunctions was recognized in this context by Ville and Bouzitat,[4] who also performed a lot of numerical work. Finally Fuchs[5] has stated, without proof, a theorem equivalent to Theorem 2. He considers n-dimensional spaces and Fourier transforms, and two arbitrary subsets of finite measure in the time- and frequency-spaces respectively. His proof, however, which we have been privileged to see, is quite different, and is not directed towards the properties of $\mathfrak{B}$ and $\mathfrak{D}$ which have been our chief concern. Our present method is capable of broad generalization; some thoughts in this direction are given in the next section.

IV. EXTENSIONS OF THE THEORY

It is quite natural for us to ask what the real essentials of the study up to this point have been, and under what circumstances results similar to Theorems 1 and 2 could be obtained. Such an investigation will be reported in a separate paper;[6] we should, however, note what some of the results are. For the relevant language, we refer the reader to Ref. 6.

We have a Hilbert space $\mathfrak{L}^2$, and two subspaces $\mathfrak{B}$ and $\mathfrak{D}$. The key property we require is that $\mathfrak{B}$ and $\mathfrak{D}$ form a nonzero minimum angle; the latter property is equivalent to requiring that

$$\sup_{f \in \mathfrak{L}^2} \frac{\| BDBf \|}{\| f \|} < 1.$$

It now follows that $\mathfrak{B} + \mathfrak{D}$ is closed, and we can again study the region of possible values of $\| Bf \|$ and $\| Df \|$ if $\| f \| = 1$. We do not, however, obtain eigenfunctions analogous to $\{\psi_n\}$ unless the operator BDB is completely continuous. If, for example, $\mathfrak{L}^2$ is the space of square-integrable functions with respect to Lebesgue measure over n-dimensional Euclidean space R^n, if $\mathfrak{D}$ is the subspace of functions vanishing outside of a bounded subset of R^n of positive measure, and if $\mathfrak{B}$ is the subspace of functions whose Fourier transforms vanish outside of another bounded subset of R^n of positive measure, then BDB is completely continuous, and the full theory applies.

As an example of a theorem which is again true in the general situation, but is of interest also for timelimited and bandlimited functions, let us prove

Theorem 3: Let an arbitrary function $d \in \mathfrak{D}$, and another function $b \in \mathfrak{B}$, be given. Then there exists an infinite collection S of functions $f \in \mathfrak{L}^2$ such that if $f \in S$, then $Df = d$ and $Bf = b$. There is a unique

$f_0 \in S$ of least energy, and there is a unique $f_1 \in S \cap (\mathfrak{B} + \mathfrak{D})$; furthermore, $f_0 = f_1$.

Proof: Let us consider the function

$$f^* = \sum_0^\infty (1 - B)(DB)^m d + \sum_0^\infty (1 - D)(BD)^m b. \tag{14}$$

The first sum, for example, means

$$d - Bd + DBd - BDBd + DBDBd - BDBDBd + \cdots .$$

Since, for any g, $\| DBg \| < \sqrt{\lambda_0} \| g \|$ and $\| BDg \| \leqq \sqrt{\lambda_0} \| g \|$, we know that the two series defined on the right side of (14) converge in norm, with their sum defined as the function $f^* \in \mathfrak{L}^2$. Furthermore, since f^* is defined as a limit of functions in $\mathfrak{B} + \mathfrak{D}$, it is, by Lemma 2, itself in $\mathfrak{B} + \mathfrak{D}$. So we may write

$$f^* = d^* + b^*,$$

where $d^* \in \mathfrak{D}$ and $b^* \in \mathfrak{B}$.

Let us next compute Df^* and Bf^*. We have

$$Df^* = \sum_0^\infty (1 - DB)(DB)^m d + \sum_0^\infty (D - D)(BD)^m b;$$

all of the second series, and all but the first half of the first term of the first series, vanish. Hence $Df^* = d$, and similarly $Bf^* = b$. We have thus shown that $f^* \in S \cap (\mathfrak{B} + \mathfrak{D})$; we can complete the proof that $f^* = f_1$ if we can show that $S \cap (\mathfrak{B} + \mathfrak{D})$ contains no other function.

Suppose that $f_i = d_i + b_i$, $i = 1, 2$ are both in $S \cap (\mathfrak{B} + \mathfrak{D})$. Then

$$d = Df_1 = d_1 + Db_1 = d_2 + Db_2 = Df_2 \tag{15}$$

and

$$b = Bf_1 = Bd_1 + b_1 = Bd_2 + b_2 = Bf_2$$

so that

$$DBd_1 + Db_1 = DBd_2 + Db_2 . \tag{16}$$

Hence, by subtracting (16) from (15), we have

$$(1 - DB)d_1 = (1 - DB)d_2 ,$$

or

$$(d_1 - d_2) = DB(d_1 - d_2).$$

Since, however, $\| DBg \| \leqq \sqrt{\lambda_0} \| g \|$ for any g, we must have $d_1 - d_2 = 0$, so that $d_1 = d_2$. Similarly, $b_1 = b_2$, so that $f_1 = f_2$, and thus f^* is the unique member of $S \cap (\mathfrak{B} + \mathfrak{D})$.

Now suppose x is any other member of S. We may write

$$x = f^* + \varphi,$$

and since $Dx = Df^* = d$ and $Bx = Bf^* = b$, it follows that

$$D\varphi = B\varphi = 0.$$

But

$$\| x \|^2 = \| f^* \|^2 + \| \varphi \|^2 + 2 \operatorname{Re}(f^*,\varphi),$$

and

$$f^* = d^* + b^* \qquad \text{while} \qquad \varphi \perp \mathfrak{D} + \mathfrak{B}.$$

Hence

$$(f^*,\varphi) = 0,$$

and

$$\| x \|^2 = \| f^* \|^2 + \| \varphi \|^2 \geqq \| f^* \|^2,$$

with equality if and only if φ vanishes. Thus f^* is also the unique member of S of minimum norm. An infinite number of other members of S may be formed by adding to f^* any of the functions orthogonal to $\mathfrak{B} + \mathfrak{D}$ whose existence is guaranteed by Lemma 3.

Note: If $d = \sum a_i D\psi_i$ and $b = \sum b_i \psi_i$, then

$$f^* = \sum \frac{a_i - b_i}{1 - \lambda_i} D\psi_i + \sum \frac{b_i - a_i\lambda_i}{1 - \lambda_i} \psi_i,$$

so that, in particular,

$$\| f^* \| \leqq \frac{1}{\sqrt{1 - \lambda_0}} (\| d \| + \| b \|).$$

V. APPLICATIONS

5.1 *Filter Theory*

Suppose we wish a filter to have an impulse response $f(t)$ which vanishes for $t > T$. Such a filter clearly cannot be strictly bandpass; but how would we select the filter so that as much of the impulse response as possible is contained in $|\omega| < \Omega$ for some given Ω? Suppose, by this, we mean to choose $f(t)$ so that

$$\frac{\int_{-\Omega}^{\Omega} | F(\omega) |^2 \, d\omega}{\int_{-\infty}^{\infty} | F(\omega) |^2 \, d\omega}$$

is as large as possible, where

$$F(\omega) = \int_0^T f(t)\, e^{-i\omega t}\, dt$$

is the Fourier transform of $f(t)$. Then the best choice is

$$f(t) = \psi_0\left(t + \frac{T}{2}, c\right),$$

where $c = \Omega T/2$, and ψ_0 is the prolate spheroidal function of the present and the preceding papers.

If, instead of requiring $f(t)$ to vanish outside of $(0,T)$, we ask that both

$$\int_{|\omega| \geqq \Omega} |F(\omega)|^2\, d\omega = \beta^2$$

and

$$\int_{-\infty}^{0} + \int_{T}^{\infty} |f(t)|^2\, dt = \alpha^2$$

be small while the total energy of the impulse response is fixed at unity; then Theorem 2 above gives the complete region of possible (α,β) values.

5.2 *Data Transmission*

When we choose a combination of pulse shape and transmission characteristic for a broadband data transmission system, we are interested in minimizing both the tail of a pulse outside its time slot and its spectrum outside of an assigned frequency band. Once again, it is not possible to make both of these "spillovers" in time and frequency arbitrarily small; the above theory gives some information on interchannel and intersymbol interference. For a theory which is more nearly complete, however, the relation between timelimiting and passbandlimiting (i.e., to $\Omega_1 \leqq |\omega| \leqq \Omega_2$) needs to be better understood; while our general results apply, the identity of the optimal function ψ_0 is not known in the case that B is projection of the transform into such a passband.

5.3 *Antenna Theory*

Let us consider a horizontal (s,t) plane from which the strip $|t| < a$ of width $2a$, to be called the *aperture*, has been removed. If the illumination across the aperture is independent of s, then the amplitude of the field across the aperture may be represented by a function $f(t)$ of one variable, where $|t| < a$. If we consider the resultant pattern of radiation in a distant parallel horizontal plane, then the field at a large distance from the aperture is proportional to

$$\int_{-a}^{a} f(t)\, e^{itu}\, dt = F(u),$$

where $u = k \sin \theta$, $k = 2\pi/\lambda$, θ is an angle measured from the vertical through the center of the aperture, and λ is the wavelength. The Q of the antenna is then defined (equivalent to the definition of Woodward and Lawson;[7] it is given explicitly by Kovács and Solymán[8]) as

$$Q = \frac{\int_{|u|>k} |F(u)|^2\, du}{\int_{-k}^{k} |F(u)|^2\, du}.$$

This may be rewritten as

$$Q = \frac{\int_{-a}^{a} |f(x)|^2\, dx}{\int_{-\infty}^{\infty} |Bf(x)|^2\, dx} - 1,$$

where B means limiting the Fourier transform of f to $|u| \leqq k$. Thus by the previous theory,

$$Q \geqq \frac{1}{\lambda_0} - 1,$$

where $\lambda_0 = \lambda_0(ak/2)$ is the first eigenvalue of (2) as defined in this and the preceding paper. We thus have an absolute lower bound on the Q which can be obtained for given a and k.

REFERENCES

1. Slepian, D. and Pollak, H. O., this issue, p. 43.
2. Chalk, J. H. H., The Optimum Pulse-Shape for Pulse Communication, Proc. I.E.E., **87**, 1950, p. 88.
3. Gurevich, M. S., Signals of Finite Duration, Containing a Maximal Part of Their Energy in a Given Bandwidth, Radiotechnika: Elektronika, **3**, 1956, p. 313.
4. Ville, J. A. and Bouzitat, J., Note sur un signal de durée finie et d'energie filtrée maximum, Cables & Trans., **11**, 1957, p. 102.
5. Fuchs, W. H. J., On the Magnitude of Fourier Transforms, Proc. Int. Math. Cong., Amsterdam, September 1954.
6. Landau, H. J. and Pollak, H. O., Subspaces Forming a Nonzero Angle, with Applications to Fourier Analysis, to be published.
7. Woodward, P. M. and Lawson, J. D., The Theoretical Precision with Which an Arbitrary Radiation Pattern May Be Obtained from a Source of Finite Size, J.I.E.E., **95**, pt. III, 1948, p. 363.
8. Kovács, R. and Solymán, L., Theory of Aperture Aerials Based on the Properties of Entire Functions of Exponential Type, Acta Phys. Budapest, **6**, 1956, p. 161.