Sampling, Data Transmission, and the Nyquist Rate

H. J. LANDAU

Abstract—**The sampling theorem for bandlimited signals of finite energy can be interpreted in two ways, associated with the names of Nyquist and Shannon.**

1) **Every signal of finite energy and bandwidth W Hz may be completely recovered, in a simple way, from a knowledge of its samples taken at the rate of $2W$ per second (Nyquist rate). Moreover, the recovery is stable, in the sense that a small error in reading sample values produces only a correspondingly small error in the recovered signal.**
2) **Every square-summable sequence of numbers may be transmitted at the rate of $2W$ per second over an ideal channel of bandwidth W Hz, by being represented as the samples of an easily constructed bandlimited signal of finite energy.**

The practical importance of these results, together with the restrictions implicit in the sampling theorem, make it natural to ask whether the above rates cannot be improved, by passing to differently chosen sampling instants, or to bandpass or multiband (rather than bandlimited) signals, or to more elaborate computations. In this paper we draw a distinction between reconstructing a signal from its samples, and doing so in a stable way, and we argue that only stable sampling is meaningful in practice. We then prove that:

1) ***stable sampling cannot be performed at a rate lower than the Nyquist,***
2) ***data cannot be transmitted as samples at a rate higher than the Nyquist,***

regardless of the location of sampling instants, the nature of the set of frequencies which the signals occupy, or the method of construction. These conclusions apply not merely to finite-energy, but also to bounded, signals.

INTRODUCTION

THE sampling theorem states that $f(t)$ is a signal of finite energy bandlimited to W Hz if and only if

$$f(t) = \sum_{-\infty}^{\infty} a_k \frac{\sin 2\pi W(t - k/2W)}{2\pi W(t - k/2W)} \quad \text{with} \quad \sum |a_k|^2 < \infty ; \quad (1)$$

thereupon

$$a_k = f(k/2W) \text{ and } \int_{-\infty}^{\infty} |f(t)|^2\, dt = \frac{1}{2W}\sum |a_k|^2 = \frac{1}{2W}\sum |f(k/2W)|^2.$$

This may be read in two ways, each of which has found important applications.

1) Every signal of finite energy and bandwidth W Hz may be completely recovered from a knowledge of its samples taken at the rate of $2W$ per second. Moreover—indispensable for any implementation in practice—the recovery is stable, in the sense that a small error in reading the sample values produces only a correspondingly small error in the recovered signal.
2) Every square-summable sequence of numbers may be transmitted at the rate of $2W$ per second over an ideal channel of bandwidth W Hz, by being represented as the samples (at the points $k/2W$) of a bandlimited signal of finite energy.

Thus the sampling theorem serves as basis for the interchangeability of analog signals and digital sequences, so valuable in modern communication systems. The rate of $2W$ samples per second for W Hz of bandwidth is often called the *Nyquist rate*.

In view of the practical importance of these results, it is natural to ask whether the Nyquist rate prescribed above for stable sampling or for data transmission cannot somehow be improved upon. Since 1) and 2) refer to a very special situation—to signals which in frequency occupy only a single band and to their values at regularly spaced instants—it is conceivable that signals might be recoverable from their values taken at a lower rate, if the sampling instants were chosen differently; or if the signals had their frequencies in a union of several bands; or at the cost of more computing than is required by the simple formula (1). Were this possible, we would have a more efficient reduction of continuous signals to digital form, with a consequent saving of bandwidth. Similarly, it is conceivable that arbitrary square-summable numbers could be prescribed at a higher rate as the sample values of a signal, if the sampling instants were properly chosen, or if the carrying signals were allowed frequencies in an appropriate union of bands. Were this possible, we would have a more efficient way of sending data over given bandwidth.

The purpose of this paper is to prove that stable sampling of signals cannot be performed at less than the Nyquist rate, nor can data be transmitted as samples at a greater rate, regardless of the location of sampling instants, the nature of the set of frequencies which the signals occupy, or the method of computation. Although discussed in the context of finite-energy signals, these results also apply to bounded signals.

FORMULATION OF THE PROBLEM AND STATEMENT OF RESULTS

We consider a fixed set S of N-frequency bands whose total finite length (counting negative as well as positive frequencies) we denote by $m(S)$, and the space $\mathcal{B}(S)$ of all signals of finite energy with frequencies contained only in S. Specifically, we make the following definition.

Manuscript received March 3, 1967; revised June 12, 1967.
The author is with Bell Telephone Laboratories, Inc., Murray Hill, N.J.

Definition: $\mathscr{B}(S)$ consists of all square-integrable functions $f(t)$ whose Fourier transform $F(\omega)$ is supported on S, i.e., $F(\omega)\equiv 0$ for $\omega \notin S$.

We have in the past called such signals multiband.

We start by considering those sets of points $t=\{t_k\}$ at which the values of a signal $f(t)\in\mathscr{B}(S)$ can be used to reconstruct the signal everywhere. We assume from the outset that these points are separated by at least some positive distance d, but otherwise do not restrict their location in any way; this separation hypothesis, while required in the proofs, does not seem a significant restriction in cases of practical interest. The first requirement to impose on the $\{t_k\}$ is that the values $\{f(t_k)\}$ determine $f(t)$, i.e., if two functions $f(t)$ and $g(t)\in\mathscr{B}(S)$ agree at the points $t=\{t_k\}$, they coincide identically. Writing this in terms of the function $h(t)=f(t)-g(t)$, we require:

$$h \in \mathscr{B}(S), \quad \text{and} \quad h(t_k) = 0 \text{ for all } k, \text{ implies } h(t) = 0. \qquad (2)$$

We will call a set $\{t_k\}$ satisfying (2) a *set of uniqueness for* $\mathscr{B}(S)$.

If $\{t_k\}$ is a set of uniqueness for $\mathscr{B}(S)$, the samples of $f\in\mathscr{B}(S)$ taken at the instants $t=\{t_k\}$ do indeed determine the signal, but this in itself does not constitute a sampling theorem adequate in practice. The reason is that we know nothing about the effect on the reconstruction process of an error in measuring the sample values. It may very well be that an arbitrarily small misreading of the numbers $f(t_k)$ will produce an arbitrarily large error in the recovered function. To guard against this, we must introduce an additional requirement of stability: small errors in reading the samples shall produce no more than correspondingly small errors in the recovered function, or, in terms of the difference between the true and reconstructed signal, if the values of a function $h\in\mathscr{B}(S)$ at the points $\{t_k\}$ are small, $h(t)$ shall itself be small. There are many ways of measuring the size of a signal; we choose energy and put our condition as follows.

Definition: The set of points $\{t_k\}$ is a *set of stable sampling for* $\mathscr{B}(S)$ if there exists a constant K such that

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt \le K \sum |f(t_k)|^2$$

for all $f(t)\in\mathscr{B}(S)$.

We can now see that a set of stable sampling is always a set of uniqueness, whereas the reverse is not true. Moreover, our definition does not commit us to any particular recovery scheme, and so describes the most general circumstances in which sampling can be performed in practice.

Taking up next the problem of transmitting data as samples, we consider instants $t=\{t_k\}$ at which square-summable values can be prescribed arbitrarily for a signal $f(t)\in\mathscr{B}(S)$. Here too we formulate the matter independently of any method of computation.

Definition: The set of points $\{t_k\}$ is a *set of interpolation for* $\mathscr{B}(S)$ if, given any set of numbers $\{a_k\}$ with $\Sigma|a_k|^2<\infty$, there exists $f(t)\in\mathscr{B}(S)$ with $f(t_k)=a_k$.

We show in the Appendix that points of a set of interpolation for $\mathscr{B}(S)$ are necessarily separated by at least some positive distance d, and that the interpolation can be performed in a stable way, i.e., the function $f(t)\in\mathscr{B}(S)$ can be chosen so that

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt \le K \sum |a_k|^2$$

with the constant K independent of $\{a_k\}$.

There is an extensive literature of particular sampling and interpolation theorems, especially when S is a single interval, but we do not intend to discuss it here. Instead, we will be concerned with the rate at which stable sampling or interpolation can be performed in $\mathscr{B}(S)$. This corresponds exactly to the density, on the t-axis, of sets of sampling or interpolation. The facts are as follows.

1) There exist sets of uniqueness for $\mathscr{B}(S)$ having arbitrarily low density[2]. This is occasionally reported as a sampling result with practical implications, but such an interpretation is somewhat misleading, in view of the lack of stability. When S is a single interval, so that $\mathscr{B}(S)$ consists of all signals of prescribed bandwidth, a remarkable result concerning sets of uniqueness is available: given a set of points $\{t_k\}$, Beurling and Malliavin[1] have shown how to determine explicitly the least upper bound of those bandwidths for which $\{t_k\}$ is a set of uniqueness.
2) If $\{t_k\}$ is a set of stable sampling for $\mathscr{B}(S)$, then every interval of length r must contain at least $(rm(S)(2\pi)^{-1} - A \log^+ r - \mathrm{B})$ of the points of $\{t_k\}$, with A and B appropriate constants.[1] Hence the density of $\{t_k\}$ is at least $m(S)/2\pi$. *Stable sampling cannot be performed at lower than the Nyquist rate.*
3) If $\{t_k\}$ is a set of interpolation for $\mathscr{B}(S)$, no interval of length r can contain more than $(rm(S)(2\pi)^{-1} + C \log^+ r + D)$ of the points of $\{t_k\}$, with C and D appropriate constants. Hence the density of $\{t_k\}$ is at most $m(S)/2\pi$. *Data cannot be sent as samples at higher than the Nyquist rate.*

We will indicate the proofs of statements 2) and 3) and close with some remarks concerning these problems for bounded, rather than finite-energy, functions.

Discussion and Proofs

We begin with an intuitive explanation of the proof. In what follows we will use the term "sampling" to mean exclusively "stable sampling."

Suppose $\{t_k\}$ is a set of sampling for $\mathscr{B}(S)$ and I is an interval containing n of the points of $\{t_k\}$. Suppose $f(t)\in\mathscr{B}(S)$ has little of its energy outside I; we will refer to such a function as being well concentrated on I. If we can presume that the samples $f(t_k)$ at points t_k outside I are also small, then, by virtue of our discussion of stability, replacing these sample values by zero should not too drastically affect our

[1] With $\log^+ r = \max(\log r, 0)$.

reconstruction of $f(t)$. We conclude that a function well concentrated on I can be substantially determined by the n measurements which give $f(t_k)$ for t_k in I, hence that the collection of all functions of $\mathscr{B}(S)$ well concentrated on I is, in some sense, no more than n-dimensional. Taken in conjunction with a similar argument about sets of interpolation, this seems to suggest that determining the number of independent signals which are well concentrated on I might yield information about our current problem. We have already had occasion to discuss the latter question in terms of the eigenvalues of a certain operator[4]. We will show that this formulation does indeed have a close connection with the sampling and interpolation which concerns us here.

With I an arbitrary given set, we begin by determining that function of $\mathscr{B}(S)$ which is most concentrated on I, i.e., that $f\in\mathscr{B}(S)$ for which the concentration

$$\lambda = \frac{\int_I |f(t)|^2\,dt}{\int_{-\infty}^{\infty} |f(t)|^2\,dt} \tag{3}$$

achieves its largest value. This is a standard problem of maximizing a quadratic form and leads immediately to the eigenvalue equation

$$\lambda_k\phi_k(t) = \frac{1}{\sqrt{2\pi}}\int_I \phi_k(x)w(t-x)\,dx, \tag{4}$$

in which the Fourier transform of w is the characteristic function of the set S. The kernel being hermitian, positive definite, and square-integrable, (4) is of the Hilbert-Schmidt type, so that eigenfunctions and eigenvalues do exist (see p. 242, Riesz and Nagy[6]); we denote the latter, in nonincreasing order, by $\lambda_k(I, S)$, $k=0, 1, 2, \cdots$, and by $\lambda_k(r, S)$ if I is a single interval of length r.

By definition, $\lambda_{j-1}(r, S)$ is the largest concentration on an interval of length r achievable by a function of $\mathscr{B}(S)$ which is orthogonal to the $j-1$ eigenfunctions $\phi_0(x), \cdots, \phi_{j-2}(x)$ previously determined; intuitively, then, $\lambda_{j-1}(r, S)$ represents the concentration of the jth most concentrated function. Our inquiry into the number of independent well concentrated functions therefore leads us to consider the behavior of the sequence $\{\lambda_k(r, S)\}$ as a function of r. There are two simple consequences of (4) which yield a lot of information:

$$\sum_{k=0}^{\infty} \lambda_k(r, S) = \text{trace} = \frac{1}{\sqrt{2\pi}}\int_0^r w(0)\,dx = \frac{m(S)r}{2\pi}, \tag{5}$$

$$\sum_{k=0}^{\infty} \lambda_k^2(r, S) = \text{double integral of } |\text{kernel}|^2. \tag{6}$$

To evaluate the latter most easily, we observe that (4) may be converted to other equations without changing the eigenvalues. One such transformation, obtained by introducing into (4) the definition of w, multiplying both sides by e^{-itu}, and integrating over I, gives

$$\lambda_k(r, S)\psi_k(t) = \frac{1}{\sqrt{2\pi}}\frac{\sqrt{2}}{\sqrt{\pi}}\int_S \psi_k(x)\frac{\sin\frac{r}{2}(t-x)}{(t-x)}\,dx, \tag{7}$$

which represents the dual problem of concentrating on S functions whose Fourier transforms are supported on an interval of length r. Now by (6)

$$\sum \lambda_k^2(r, S) = \frac{1}{\pi^2}\int_S\int_S \left|\frac{\sin\frac{r}{2}(t-x)}{(t-x)}\right|^2 dx\,dt.$$

If S consists of intervals of length l_i, we find, after some manipulation,

$$\begin{aligned}\sum \lambda_k^2(r, S) &\geq \sum_{i=1}^{N}\frac{1}{\pi^2}\int_0^{l_i}\int_0^{l_i}\frac{\sin^2\frac{r}{2}(t-x)}{(t-x)^2}\,dx\,dt\\ &\geq \sum_{i=1}^{N}\left(\frac{l_i r}{2\pi} - \frac{1}{\pi^2}\log^+ l_i r - 1\right)\\ &\geq \frac{m(S)r}{2\pi} - A\log^+ r - B,\end{aligned} \tag{8}$$

with constants A and B depending only on S. Thus, from (5) and (8),

$$\sum_k \lambda_k(r, S)\{1-\lambda_k(r, S)\} \leq A\log^+ r + B. \tag{9}$$

It follows from (9) that the number of eigenvalues $\lambda_k(r, S)$ lying between any two fixed bounds $\delta>0$ and $\gamma<1$ can grow at most logarithmically with r. Specifically, each such eigenvalue contributes to the left-hand side of (9) an amount no smaller than $\alpha=\min\{\delta(1-\delta), \gamma(1-\gamma)\}>0$; hence the number of these eigenvalues cannot exceed $A/\alpha\log^+ r+B/\alpha$. Taken together with (5) and (8), this observation shows that, asymptotically in r, the number of $\lambda_k(r, S)$ near 1 behaves like $m(S)(2\pi)^{-1}r$. But we can sharpen this statement by means of the following two lemmas (proved in the Appendix), which serve as well to establish the connection between the eigenvalue problem and sampling and interpolation.

Lemma 1

Let S be bounded and $\{t_k\}$ be a set of sampling for $\mathscr{B}(S)$, whose points are separated by at least $d>0$. Let I be any compact set, $I+$ be the set of points whose distance to I is less than $d/2$, and $n(I+)$ be the number of points of $\{t_k\}$ contained in $I+$. Then $\lambda_{n(I+)}(I, S)\leq\gamma<1$, where γ depends on S and $\{t_k\}$ but not on I.

Lemma 2

Let S be bounded and $\{t_k\}$ be a set of interpolation for $\mathscr{B}(S)$, whose points are separated by at least $d>0$. Let I be any compact set, $I-$ be the set of points whose distance to the complement of I exceeds $d/2$, and $n(I-)$ be the number of points of $\{t_k\}$ contained in $I-$. Then $\lambda_{n(I-)-1}(r, S)\geq\delta>0$, where δ depends on S and $\{t_k\}$ but not on I.

We have already remarked that the eigenvalues $\lambda_k(I, S)$ remain unchanged when we reverse the roles of the time and frequency domains. It is also easy to see from (4) that they are unaffected by a rescaling of time and frequency which replaces I by αI and S by $\alpha^{-1}S$, with α any constant.[2] By applying both of these permissible transformations, we conclude that $\lambda_{k-1}(r, S)$ is also the kth eigenvalue of the problem of concentrating functions whose frequencies lie in a single interval of length 2π (i.e., bandlimited functions) on the set $(r/2\pi)S$. But now the ordinary sampling theorem shows the integer points to be a set both of sampling and interpolation. When S is the union of N intervals, the number of these points contained in $S+$ is at most $(2\pi)^{-1}rm(S)+2N$, and their number in $S-$ is at least $(2\pi)^{-1}rm(S)-2N$. Thus Lemmas 1 and 2 establish the existence of absolute constants γ_0 and δ_0 such that

$$\lambda_{[(2\pi)^{-1}rm(S)]+2N}(r, S) \leq \gamma_0 < 1, \tag{10}$$

$$\lambda_{[(2\pi)^{-1}rm(S)]-2N-1}(r, S) \geq \delta_0 > 0; \tag{11}$$

here we use the notation $[x]$ to denote the integer part of x. It follows from their construction that γ_0 and δ_0 may be taken to be $\frac{1}{2}$, but we will not prove this here.

In the preceding argument we used Lemmas 1 and 2, together with known sets of sampling and interpolation, to determine the behavior of the $\{\lambda_k(r, S)\}$. Now the lemmas allow us in turn to apply this information to arbitrary sets of sampling or interpolation.

Theorem 1

If $\{t_k\}$ is a set of sampling for $\mathscr{B}(S)$, every interval of length r must contain at least $(rm(S)(2\pi)^{-1}-A\log^+ r-B)$ of the points of $\{t_k\}$, with A and B appropriate constants.

Proof: Let I be an interval of length r, and $n(I)$ the number of points of $\{t_k\}$ contained in I; since I is a single interval, $n(I+)\leq n(I)+2$. By Lemma 1, there exists a constant $\gamma<1$ independent of r such that

$$\lambda_{n(I)+2}(r, S) \leq \gamma < 1. \tag{12}$$

But from (11)

$$\lambda_{[rm(S)(2\pi)^{-1}]-2N-1}(r, S) \geq \delta_0 > 0, \tag{13}$$

and, as we have argued in connection with (9), the number of eigenvalues lying in the range $\delta_0<\lambda_k(r, S)\leq\gamma$ increases at most logarithmically with r. Thus

$$([rm(S)(2\pi)^{-1}] - 2N - 1) - (n(I) + 2) \leq A' \log^+ r + B',$$

with constants A' and B', whence

$$n(I) \geq rm(S)(2\pi)^{-1} - A\log^+ r - B.$$

Theorem 1 is established.

It follows immediately that the density of a set of sampling cannot be lower than $m(S)/2\pi$.

Theorem 2

If $\{t_k\}$ is a set of interpolation for $\mathscr{B}(S)$, no interval of length r can contain more than $(rm(S)(2\pi)^{-1}+C\log^+ r+D)$ of the points of $\{t_k\}$, with C and D appropriate constants.

[2] We denote by αI the set of points of the form αt for $t\in I$.

Proof: We follow the proof of Theorem 1, interchanging the roles of Lemmas 1 and 2. Thus with I an interval of length r, $n(I-)\geq n(I)-2$, and, by Lemma 2, there exists a constant $\delta>0$ independent of r such that

$$\lambda_{n(I)-3}(r, S) \geq \delta > 0. \tag{14}$$

But from (10)

$$\lambda_{[rm(S)(2\pi)^{-1}]+2N} \leq \gamma_0 < 1,$$

so that

$$(n(I) - 3) - ([rm(S)(2\pi)^{-1}] + 2N) \leq C' \log^+ r + D',$$

whence

$$n(I) \leq rm(S)(2\pi)^{-1} + C\log^+ r + D.$$

Theorem 2 is established.

It follows immediately that the density of a set of interpolation cannot be higher than $m(S)/2\pi$.

Concluding Remarks

1) The aforementioned density results apply equally well to functions of several variables and to frequency sets S more complicated than finite unions of intervals. A complete account of this extension is to be found elsewhere[3].

2) It may well be argued that functions of finite energy are too special a class, and that bounded functions present greater practical interest. The problem of describing frequency content is rather severe in this case, the Fourier transform not being available, but meaningful definitions can be given (using distributions, for example) for the class $\mathscr{B}_b(S)$ of bounded signals with frequencies in S; we will not discuss the details here. We can then formulate the notions of stable sampling and interpolation as follows.

Definition: The points $\{t_k\}$ form a set of stable sampling for $\mathscr{B}_b(S)$ if there exists a constant K such that

$$\sup_t |\phi(t)| \leq K \sup_k |\phi(t_k)|,$$

for all $\phi\in\mathscr{B}_b(S)$.

Definition: The points $\{t_k\}$ form a set of interpolation for $\mathscr{B}_b(S)$ if, given any bounded sequence of numbers $\{a_k\}$, there exists a function $\phi(t)\in\mathscr{B}_b(S)$ with $\phi(t_k)=a_k$.

Such sets were first defined and studied by Beurling. One can show the following.

a) The density of a set of sampling for $\mathscr{B}_b(S)$ is always strictly larger than $m(S)/2\pi$; thus, for bounded signals, stable sampling requires a greater than Nyquist rate.
b) The density of a set of interpolation for $\mathscr{B}_b(S)$ is at most $m(S)/2\pi$, and strictly smaller if S is a single interval. We suspect it to be always strictly smaller, but have not succeeded in proving this. In any case, transmission of data as samples of bounded signals can certainly not be performed at a greater than Nyquist rate.

Details may again be found elsewhere[3].

Appendix

Proposition 1

Let S be bounded, and let $\{t_k\}$ be a set of interpolation for $\mathscr{B}(S)$. Then the points of $\{t_k\}$ are separated by at least some positive distance d, and the interpolation can be performed in a stable way.

Proof: By writing $f \in \mathscr{B}(S)$ as the inverse transform of its Fourier transform, and applying Schwarz's inequality and Parseval's theorem to this representation, we find

$$|f(t)|^2 \le K_1 \int_{-\infty}^{\infty} |f(t)|^2 \, dt, \tag{15}$$

$$|f'(t)|^2 \le K_2 \int_{-\infty}^{\infty} |f(t)|^2 \, dt, \tag{16}$$

with constants K_1 and K_2 depending only on S. Thus for signals in $\mathscr{B}(S)$ a bound on energy implies a corresponding bound on both the signal and its derivative.

If the points $\{t_k\}$ are not separated by at least some fixed positive distance, we can find two disjoint subsequences, which we denote by $\{p_k\}$ and $\{q_k\}$, with the property that $|p_k - q_k| \le 1/k^2$, $k = 1, 2, \cdots$. Now let us prescribe the values $1/k$ at $t = p_k$, $-1/k$ at $t = q_k$, and 0 at every remaining point of $\{t_k\}$; since these values are square-summable, and $\{t_k\}$ is a set of interpolation, they will be assumed by some $f(t) \in \mathscr{B}(S)$. But then $2/k = f(p_k) - f(q_k) \le |p_k - q_k| \, |f'(s_k)|$ with s_k a point lying between p_k and q_k. Hence $|f'(s_k)| \ge 2k$, $k = 1, 2, \cdots$, contradicting (16). This establishes the first part of the proposition.

To prove the second part, we appeal to an abstract principle. In $\mathscr{B}(S)$, let us view functions as vectors and energy as the square of length; this makes $\mathscr{B}(S)$ into a Hilbert space. Let $\mathscr{E}^0(S)$ be the subspace of $\mathscr{B}(S)$ consisting of those functions which vanish at all the points $\{t_k\}$. This subspace is closed, since, by (15), convergence in energy of functions of $\mathscr{E}^0(S)$ to a function $h(t)$ implies uniform convergence as well, so that $h(t)$ also must vanish on $\{t_k\}$. The reason for introducing $\mathscr{E}^0(S)$ is that it offers the possibility of reducing the energy of interpolating functions. Specifically, starting with one $f(t)$ in $\mathscr{B}(S)$ which interpolates to a prescribed set of values $\{a_k\}$ at $\{t_k\}$, we can form the collection $\{f(t) + g(t)\}$ with $g(t) \in \mathscr{E}^0(S)$; all of these functions continue to interpolate, and we may ask for that member of the collection which has least energy or length. This problem has as its solution the function which represents the projection of f onto the orthogonal complement in $\mathscr{B}(S)$ of $\mathscr{E}^0(S)$, a subspace which we denote by $\mathscr{E}(S)$. Thus interpolation can be performed in $\mathscr{E}(S)$, whereupon the interpolating function is uniquely determined and has minimum energy.

Now let us consider the transformation T which assigns to a square-summable sequence $\{a_k\}$ that function of $\mathscr{E}(S)$ which interpolates to a_k at $t = t_k$. By hypothesis, T is defined on the whole space l^2 of square-summable sequences, maps l^2 into the Hilbert space $\mathscr{E}(S)$, and, in view of (15), has a closed graph. Then by the closed graph theorem (see Loomis[5], p. 18) T is bounded, i.e., there exists a constant K such that the interpolating functions $f(t)$ in $\mathscr{E}(S)$ satisfy

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt \le K \sum |a_k|^2 = K \sum |f(t_k)|^2.$$

This shows that when interpolation is performed in $\mathscr{E}(S)$, it is necessarily stable. Proposition 1 is established.

Lemma 1

Let S be bounded and $\{t_k\}$ be a set of sampling for $\mathscr{B}(S)$, whose points are separated by at least $d > 0$. Let I be any compact set, $I+$ be the set of points whose distance to I is less than $d/2$, and $n(I+)$ be the number of points of $\{t_k\}$ contained in $I+$. Then $\lambda_{n(I+)}(I, S) \le \gamma < 1$, where γ depends on S and $\{t_k\}$ but not on I.

Proof: By the Weyl-Courant lemma,

$$\lambda_k(I, S) \le \sup_{\substack{f \in B(S) \\ f \perp C_k}} \frac{\int_I |f(t)|^2 \, dt}{\int_{-\infty}^{\infty} |f(t)|^2 \, dt}, \tag{17}$$

with C_k any subspace of dimension k. Thus to prove Lemma 1, it is sufficient to show that we can keep the concentration of a function of $\mathscr{B}(S)$ over I bounded away from 1 by imposing $n(I+)$ orthogonality conditions. The following simple device accomplishes this. Let $h(y)$ be a square-integrable function which vanishes for $|y| > d/2$ and whose Fourier transform $H(\omega)$ satisfies $|H(\omega)| \ge 1$ for $\omega \in S$. Such a function certainly exists, since the transforms of functions vanishing on $|y| > d/2$ can be made to approximate uniformly an arbitrary continuous function on any bounded set. Now given $f \in \mathscr{B}(S)$, we form

$$\begin{aligned} g(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y) h(t - y) \, dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{|t-y| < d/2} f(y) h(t - y) \, dy, \end{aligned} \tag{18}$$

whence by Schwarz's inequality

$$|g(t)|^2 \le c \int_{|t-y| < d/2} |f(y)|^2 \, dy, \tag{19}$$

with $c = (2\pi)^{-1} \int |h(t)|^2 \, dt$. We observe that the Fourier transforms F and G of f and g, respectively, are related by $G = FH$. It follows that $g \in \mathscr{B}(S)$, so that by the definition of $\{t_k\}$

$$\int_{-\infty}^{\infty} |g(t)|^2 \, dt \le K \sum |g(t_k)|^2, \tag{20}$$

and by Parseval's theorem

$$\begin{aligned} \int_{-\infty}^{\infty} |f(t)|^2 \, dt &= \int_S |F(\omega)|^2 \, d\omega \le \int_S |G(\omega)|^2 \, d\omega \\ &= \int_{-\infty}^{\infty} |g(t)|^2 \, dt. \end{aligned} \tag{21}$$

Now subjecting f to the $n(I+)$ orthogonality conditions

$$\int_{-\infty}^{\infty} f(y)h(t_k - y)\,dy = 0, \qquad t_k \in I+,$$

ensures that $g(t_k)=0$, $t_k \in I+$, whereupon the sum in (20) is extended only over $t_k \notin I+$. For those t_k, the corresponding integrals in (19) are taken over non-overlapping intervals, each contained outside I; thus combining (21), (20), and (19) we find

$$\int_{-\infty}^{\infty} |f(t)|^2\,dt \le Kc \int_{t \notin I} |f(t)|^2\,dt$$

or

$$\frac{\int_I |f(t)|^2\,dt}{\int_{-\infty}^{\infty} |f(t)|^2\,dt} \le 1 - \frac{1}{Kc} = \gamma < 1.$$

The constant γ depends on S and $\{t_k\}$ since c and K do, but not on I. Applied to (17), this proves Lemma 1.

Lemma 2

Let S be a bounded set and $\{t_k\}$ be a set of interpolation for $\mathscr{B}(S)$, whose points are separated by at least $d>0$. Let I be any compact set, $I-$ be the set of points whose distance to the complement of I exceeds $d/2$, and $n(I-)$ be the number of points of $\{t_k\}$ contained in $I-$. Then $\lambda_{n(I-)-1}(I, S) \ge \delta > 0$, where δ depends on S and $\{t_k\}$ but not on I.

Proof: We again begin with the Weyl-Courant lemma in the form

$$\lambda_{k-1}(I, S) \ge \inf_{\substack{f\in\mathscr{B}(S)\\ f\in C_k}} \frac{\int_I |f(t)|^2\,dt}{\int_{-\infty}^{\infty} |f(t)|^2\,dt}, \tag{22}$$

with C_k any subspace of dimension k. Any proof of (17) can be modified to yield (22), but the following independent argument serves as well. Let L_{k-1} be the subspace of $\mathscr{B}(S)$ spanned by $\phi_0, \cdots, \phi_{k-2}$, the first $k-1$ eigenfunctions. Since the dimension of L_{k-1} is lower than that of C_k, any linear transformation of C_k into L_{k-1}, in particular the orthogonal projection onto L_{k-1}, must annihilate some nonzero element of C_k. Thus C_k contains a function $f \not\equiv 0$ orthogonal to L_{k-1}, and (22) follows from the definition of λ_{k-1}.

We have shown in Proposition 1 that it is possible to perform the interpolation in a subspace $\mathscr{E}(S)$ of $\mathscr{B}(S)$, in which every function $g(t)$ satisfies

$$\int_{-\infty}^{\infty} |g(t)|^2\,dt \le K \sum |g(t_k)|^2, \tag{23}$$

with some fixed constant K. For each t_k, we let $\phi_k(t) \in \mathscr{E}(S)$ be the function whose value is 1 at t_k and 0 at every t_j, $j \ne k$; these functions are linearly independent.

We let h be the same as in the proof of Lemma 1, and construct functions $\psi_k \in \mathscr{B}(S)$ such that

$$\phi_k(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_k(y)h(t-y)\,dy; \tag{24}$$

to see that these exist, we need only take the Fourier transform of (24) and note that $|H(\omega)| \ge 1$ for $\omega \in S$. The ψ's are likewise linearly independent, so that, letting C be the subspace of $\mathscr{B}(S)$ spanned by the functions ψ_k with t_k in $I-$, we obtain a subspace of dimension $n(I-)$.

Now given $f \in C$, we form the function g of (18). In view of (24), g is a linear combination of the ϕ_k, hence lies in $\mathscr{E}(S)$ and satisfies (21). However, since the only ϕ_k figuring in g are those with t_k in $I-$, the values of g necessarily vanish at every other t_k, so that

$$\int_{-\infty}^{\infty} |g(t)|^2\,dt \le K \sum_{I-} |g(t_k)|^2. \tag{25}$$

The relations (19) and (21) between f and g continue to hold, and when the points t_k are in $I-$ the corresponding integrals in (19) are extended over non-overlapping intervals, each contained in I, so that

$$\sum_{I-} |g(t_k)|^2 \le c \int_I |f(t)|^2\,dt. \tag{26}$$

Combining (26), (25), and (21) shows

$$\frac{\int_I |f(t)|^2\,dt}{\int_{-\infty}^{\infty} |f(t)|^2\,dt} \ge \frac{1}{Kc} = \delta > 0$$

which, applied to (22), proves Lemma 2. As in Lemma 1, the constant δ does not depend on I.

References

[1] A. Beurling and P. Malliavin, "On the closure of characters and zeros of entire functions," *Acta Mathematica*, vol. 118, pp. 79–93, May 1967.

[2] H. J. Landau, "A sparse regular sequence of exponentials closed on large sets," *Bull. Am. Math. Soc.*, vol. 70, pp. 566–569, July 1964.

[3] ——, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, pp. 37–52, February 1967.

[4] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty. III," *Bell Sys. Tech. J.*, vol. 41, pp. 1295–1336, July 1962.

[5] L. H. Loomis, *An Introduction to Abstract Harmonic Analysis*. New York: Van Nostrand, 1953.

[6] F. Riesz and B. S. Nagy, *Functional Analysis*. New York: Ungar, 1955.