

# Spatial Concentration of Wave-Fields: Towards Spatial Information Content in Arbitrary Multipath Scattering

Rodney A. Kennedy and Thushara D. Abhayapala

**Abstract**—The fundamental limitation of using a spatial channel as an information bearing resource is considered. Such theoretical foundations are largely absent in the development of wireless communication systems which unduly focus on implementation and as a consequence there are a number of misconceptions and mal-implementations of communication systems which employ the MIMO principle. The notion of essential dimensionality of spatial systems is examined in detail for the case of a spherical region in three dimensional space in a nearfield scattering environment as a pointer to a more general theory nearing completion. A non-trivial but tight analogy is drawn with the classical work on time-frequency concentration and the dimensionality of essentially time- and band-limited signals by Slepian, Landau and Pollack.

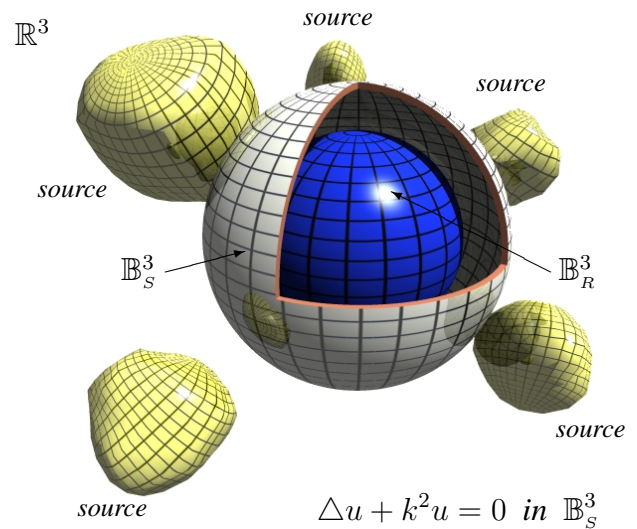
**Index Terms**—Essential dimension of space, space-time communication limits, time-frequency concentration.

## I. INTRODUCTION

A central result in information theory relates to time-frequency concentration and the essential dimensionality of time-frequency signals governed by the Fourier Transform [1–3]. When constrained in both time and frequency there is a limit to the degree of concentration, as measure by fraction-out-of-band, of energy simultaneously possible in the two domains (FOBE). This is a form of uncertainty principle where the criterion for time-frequency concentration significantly differs from the classical Heisenberg formulation which expresses signal concentration in terms of root mean square deviation (RMS). This paper deals with a generalization where one can investigate, for the first time, the fundamental limits to the use of free-space as an information bearing resource a central concept for wireless communication. The work has an impact on the theoretical limits of MIMO systems using multiple antennas by revealing a number of misconceptions related to the importance of the number of antennas — for example, the dependence of capacity of multi-antenna systems being attributed to the number of antennas will be seen to be an artifact of implementation (strictly an infinitely dispersed sparse antenna array) and not fundamentally relate to the theoretical limit of such systems. The latter can be developed using the ideas in this paper.

The notion of the essential dimension, despite the manner in which the results were originally presented, does not crucially hinge on the Fourier Transform — it is a misconception

The authors are with National ICT Australia, and the Department of Telecommunications Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra ACT 0200, Australia, e-mail:{rodney.kennedy, thushara.abhayapala}@anu.edu.au.



**Fig. 1:** Source-field and Wave-field Geometry: The region of interest is the inner ball of radius  $R$ ,  $\mathbb{B}_R^3$ . Any source for the wave-field, symbolized by the irregular outer blobs, are outside an outer ball of radius  $S$ ,  $\mathbb{B}_S^3$ , shown with cutaway. That is, the source-field is non-zero only in  $\mathbb{R}^3 \setminus \mathbb{B}_S^3$  and  $R < S$ .

to regard the Fourier Transform as playing a *necessary* role in any uncertainty principle (RMS, FOBE, or otherwise). Uncertainty principles and essential dimensionality results are a natural consequences of broad classes of operator equations where the operator is either self-adjoint, normal or unitary (and possibly more general types of operators) [4].

Wireless communications involves the exploitation of space to achieve communication. The extent to which this is fundamentally possible is constrained by the wave equation in free space. That is, the degree to which data, in an abstract sense, can be borne on information bearing wave-fields in space is limited by the essential dimensionality of such wave-fields — a concept which is developed in this paper and extends earlier work [5, 6]. The results for establishing the essential dimensionality of space have strong analogies with the seminal work on essential dimensionality of time-frequency signals [1–3]. In this work we develop the results for narrowband and in this regard frequency plays no role nor does the Fourier Transform.

## II. WAVES IN BALLS AND BLOBS

### A. Homogeneous Helmholtz Equation

The homogeneous Helmholtz equation describing a *wave-field*  $u(\mathbf{x})$  in space  $\mathbf{x} \in \mathbb{R}^3$ , also referred to as the reduced

wave equation, is given by

$$\Delta u(\mathbf{x}) + k^2 u(\mathbf{x}) = 0, \quad (1)$$

where  $\Delta$  is the Laplacian, and  $k$  is the wave number given by the real positive constant  $k = 2\pi/\lambda$  with  $\lambda$  the wavelength. Equation (1) holds in any region of space, a subset of  $\mathbb{R}^3$ , that *excludes any sources*. The formulation explicitly caters for nearfield sources by choosing the region as small as necessary. In this work we first consider regions which are closed balls, that is, spheres and their interiors:

$$\mathbb{B}_T^3 \triangleq \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| \leq T\}. \quad (2)$$

for various radii  $T$  where  $|\cdot|$  denotes euclidean distance. The unit sphere is, following convention, denoted  $\mathbb{S}^2 \triangleq \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = 1\}$ . Later in the paper, balls are generalized to blobs (to be defined) and a more abstract and general theory is developed.

The complete Helmholtz equation takes the form

$$\Delta u(\mathbf{x}) + k^2 u(\mathbf{x}) = f(\mathbf{x}), \quad (3)$$

where, in our terminology,  $f(\mathbf{x})$  is the *source-field*. The broad interest in this work is to determine the relationship between the wave-field  $u(\mathbf{x})$  and the source-field  $f(\mathbf{x})$  as governed by (3) that leads to the homogeneous equation (1) being satisfying in some user prescribed “region of interest” which is a subset of  $\mathbb{R}^3$ . In engineering terms, this region of interest is the spatial domain where antennas are located and would almost without exception exclude any sources. However such sources, carrying the signals to be detected, can be arbitrarily close to the region of interest and hence in the nearfield. Therefore, the theory to be developed is powerful enough to deal with nearfield, farfield, point and distributed sources. That is, any practical multipath field.

### B. Hilbert Space Representation

We assume familiarity with separable Hilbert Spaces with the associated concepts of complete orthonormal sequences, orthogonality, inner products, projection, Parseval relation, strong convergence, generalized Fourier representations which we use without proof [7, 8].

All solutions to (1) for a given source-free region define a linear subspace of functions which follows from the linearity and homogeneity of (1). That is, if  $u_1(\mathbf{x})$  and  $u_2(\mathbf{x})$  are solutions to (1) in a region then  $\xi_1 u_1(\mathbf{x}) + \xi_2 u_2(\mathbf{x})$  is also a solution in the same region, where  $\xi_1, \xi_2 \in \mathbb{C}$ . We formalize and develop this interpretation as follows.

We begin with the complex separable Hilbert Space of complex valued square integrable functions,  $L^2(\mathbb{B}_S^3)$ , defined in  $\mathbb{B}_S^3$  for some radius  $S$  (which later will be identified with any radius that excludes all sources), equipped with inner product

$$\langle f, g \rangle_{\mathbb{B}_S^3} \triangleq \int_{\mathbb{B}_S^3} f(\mathbf{x}) \overline{g(\mathbf{x})} dv(\mathbf{x}), \quad (4a)$$

$$\equiv \int_0^S \int_0^\pi \int_0^{2\pi} f(r, \theta, \phi) \overline{g(r, \theta, \phi)} r^2 \sin \theta d\phi d\theta dr, \quad (4b)$$

with induced norm

$$\|f\|_{\mathbb{B}_S^3}^2 \triangleq \int_{\mathbb{B}_S^3} |f(\mathbf{x})|^2 dv(\mathbf{x}), \quad (5)$$

where the volume element is given by

$$dv(\mathbf{x}) \triangleq r^2 \sin \theta d\phi d\theta dr, \quad (6)$$

and where  $(r, \theta, \phi) \equiv \mathbf{x}$  are spherical coordinates. In coordinate free terms, we can use  $|\mathbf{x}|$  for  $r$  and  $\hat{\mathbf{x}} \triangleq \mathbf{x}/|\mathbf{x}|$  for direction  $(\theta, \phi)$ . Similarly, on the *unit* sphere,  $\mathbb{S}^2$ , we use the notation

$$ds(\hat{\mathbf{x}}) \triangleq \sin \theta d\phi d\theta, \quad (7)$$

to denote a surface element. Finally, we also use the identity

$$dv(\mathbf{x}) \triangleq r^2 dr ds(\hat{\mathbf{x}}), \quad (8)$$

which combines (6) and (7).

The solutions to (1) that form a strict linear subspace of  $L^2(\mathbb{B}_S^3)$  is a separable Hilbert Space and denoted

$$\mathcal{G}_S \triangleq \{u \in L^2(\mathbb{B}_S^3) : \Delta u + k^2 u = 0\}. \quad (9)$$

Both  $L^2(\mathbb{B}_S^3)$  and its subspace  $\mathcal{G}_S$  have inner product (4).<sup>1</sup>

### C. Complete Orthonormal Sequences

Since  $\mathcal{G}_S$  is separable, there exist orthonormal sequences,  $\{\varphi_p\}_{p=0}^\infty$ , which are complete in  $\mathcal{G}_S$ , that is, for all  $u \in \mathcal{G}_S$

$$u = \sum_{p=0}^\infty \langle u, \varphi_p \rangle_{\mathbb{B}_S^3} \varphi_p, \quad (10)$$

where convergence is in the mean (strong convergence in the norm) and

$$\beta_p \triangleq \langle u, \varphi_p \rangle_{\mathbb{B}_S^3} \quad (11)$$

are the Fourier Coefficients. By completeness we have Parseval's Relation

$$\|u\|_{\mathbb{B}_S^3}^2 = \|\beta\|_{\ell^2}^2. \quad (12)$$

### D. Normalized Modes as Orthonormal Sequences

In this paper, we make use of a special complete orthonormal sequence, based on the solution to (1) by separation of variables in spherical coordinates, see [9]. This sequence is built from the modes, which form elementary solutions to (1), and take the form

$$u_n^m(\mathbf{x}) \triangleq i^n j_n(k|\mathbf{x}|) Y_n^m(\hat{\mathbf{x}}) \quad (13)$$

<sup>1</sup>The actual space is actually more restricted than indicated as we require the wave-field to be sufficiently differentiable for the Helmholtz equation to be well-defined.

where the indices are  $n = 0, 1, 2, \dots$  and  $m = -n, \dots, n$ .<sup>2</sup> Here  $j_n(k|\mathbf{x}|)$  is the spherical Bessel function of integer order  $n$ ,  $Y_n^m(\hat{\mathbf{x}}) \equiv Y_n^m(\theta, \phi)$  are the spherical harmonics

$$Y_n^m(\theta, \phi) \triangleq \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos \theta) e^{im\phi}, \quad (14)$$

which are orthonormal on the unit sphere  $\mathbb{S}^2$  with respect to the inner product

$$\langle f, g \rangle_{\mathbb{S}^2} \triangleq \int_{\mathbb{S}^2} f(\hat{\mathbf{x}}) \overline{g(\hat{\mathbf{x}})} ds(\hat{\mathbf{x}}), \quad (15a)$$

$$\equiv \int_0^\pi \int_0^{2\pi} f(\theta, \phi) \overline{g(\theta, \phi)} \sin \theta d\phi d\theta, \quad (15b)$$

and  $P_n^m(\cdot)$  are the associated Legendre functions. For  $m = 0$ , these orthogonal functions reduce to the Legendre Polynomials denoted  $P_n(\cdot)$ .

The modes, (13), have a “poly-orthogonality” meaning that they are *orthogonal for any spherically symmetric region* — a property that is analogous to the multiple orthogonality of the Prolate Spheroidal Wave Functions used in the time-frequency dimensionality case [1]. This follows mathematically from the orthonormality of the spherical harmonics and physically from the modes, (13), being obtained through separation of variables. That is, the poly-orthogonality can be expressed through

$$\int_{\mathbb{S}^2} u_n^m(|\mathbf{x}|, \hat{\mathbf{x}}) \overline{u_q^p(|\mathbf{x}|, \hat{\mathbf{x}})} ds(\hat{\mathbf{x}}) = [j_n(k|\mathbf{x}|)]^2 \delta_{mp} \delta_{nq}. \quad (16)$$

This implies orthogonality with respect to a wide selection of inner products defined on  $\mathbb{B}_S^3$  not just (4). With  $\langle u_n^m, u_q^p \rangle$  denoting a general inner product, (16) implies

$$\langle u_n^m, u_q^p \rangle = \int_{\mathbb{B}_S^3} h(|\mathbf{x}|) u_n^m(|\mathbf{x}|, \hat{\mathbf{x}}) \overline{u_q^p(|\mathbf{x}|, \hat{\mathbf{x}})} dv(\mathbf{x}) \quad (17a)$$

$$= \int_0^S h(r) [j_n(kr)]^2 r^2 dr \delta_{mp} \delta_{nq} \quad (17b)$$

where the only condition on  $h(|\mathbf{x}|)$  is that it is non-negative and has appropriate support to yield a valid inner product.

It is a simple matter to normalize the modes (13) to any spherically symmetric region. Hence for  $\mathbb{B}_S^3$ , normalization with respect to inner product (4), yields

$$\{\varphi_{n;s}^m(\mathbf{x})\}_{m,n} \triangleq \left\{ \frac{i^n j_n(k|\mathbf{x}|) Y_n^m(\hat{\mathbf{x}})}{(\int_0^S [j_n(kr)]^2 r^2 dr)^{1/2}} \right\}_{m,n}. \quad (18)$$

Based on the the denominator in (18) and because of its importance later, we define

$$\mathcal{J}_n(T) \triangleq \int_0^T [j_n(kr)]^2 r^2 dr \quad (19)$$

<sup>2</sup>It is clear that the indices  $m$  and  $n$  can be put into a bijection with the countable index set  $p$ , as in (10). Hence the modes, (13), are analogous to the  $\varphi_p$  in (10) apart from an absence of normalization. Further, in defining the modes we have included the rogue factor  $i^n$  which is standard to align with some key but esoteric identities.

which is a monotonically increasing positive function which can be expressed in closed form and grows like  $T/(2k^2)$  for large  $T$ . Define the shorthand

$$\sum_{m,n} \triangleq \sum_{n=0}^{\infty} \sum_{m=-n}^n \quad (20a)$$

$$\sum_{m,n \leq N} \equiv \sum_{n=0}^N \sum_{m=-n}^n \quad \text{and} \quad \sum_{m,n > N} \triangleq \sum_{n=N+1}^{\infty} \sum_{m=-n}^n. \quad (20b)$$

Using (18), any  $u \in \mathcal{G}_S$  has the orthonormal expansion

$$u(\mathbf{x}) = \sum_{m,n} \beta_{n;s}^m \varphi_{n;s}^m(\mathbf{x}) \quad (21a)$$

$$= \sum_{m,n} \beta_{n;s}^m \frac{i^n j_n(k|\mathbf{x}|) Y_n^m(\hat{\mathbf{x}})}{[\mathcal{J}_n(S)]^{1/2}} \quad (21b)$$

with Fourier coefficients

$$\beta_{n;s}^m \triangleq \langle u(\mathbf{x}), \varphi_{n;s}^m(\mathbf{x}) \rangle_{\mathbb{B}_S^3} \quad (22a)$$

$$= \frac{(-i)^n}{[\mathcal{J}_n(S)]^{1/2}} \int_{\mathbb{B}_S^3} u(\mathbf{x}) j_n(k|\mathbf{x}|) \overline{Y_n^m(\hat{\mathbf{x}})} dv(\mathbf{x}) \quad (22b)$$

which satisfy Parseval's Relation (12). Some examples of (21) will now be given.

### E. Examples

EXAMPLE (PLANE WAVES). Using the Jacobi-Anger Expansion [9, p.32] we have

$$e^{ik\mathbf{x} \cdot \hat{\mathbf{y}}} = \sum_{n=0}^{\infty} i^n (2n+1) j_n(k|\mathbf{x}|) P_n(\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}) \quad (23a)$$

$$= 4\pi \sum_{m,n} i^n j_n(k|\mathbf{x}|) Y_n^m(\hat{\mathbf{x}}) \overline{Y_n^m(\hat{\mathbf{y}})} \quad (23b)$$

where we have used the Addition Formula for the Spherical Harmonics [9, p.27] to go from (23a) to (23b). The left hand side is a plane wave with direction of propagation  $\hat{\mathbf{y}}$  and is a basic solution to (1) in  $\mathcal{G}_S$ . That is, for the plane wave, we have the Fourier Coefficients

$$\beta_{n;s}^m = 4\pi [\mathcal{J}_n(S)]^{1/2} \overline{Y_n^m(\hat{\mathbf{y}})} \quad (24)$$

which depend only on direction of propagation  $\hat{\mathbf{y}}$ .

EXAMPLE (POINT SOURCES). The fundamental solution to the Helmholtz equation, a point source at  $\mathbf{y}$ , is given by [9, p.5]

$$\Phi(\mathbf{x}, \mathbf{y}) \triangleq \frac{ik}{4\pi} h_0^{(1)}(k|\mathbf{x} - \mathbf{y}|) \equiv \frac{e^{ik|\mathbf{x} - \mathbf{y}|}}{4\pi|\mathbf{x} - \mathbf{y}|}, \quad \mathbf{x} \neq \mathbf{y} \quad (25)$$

where  $h_n^{(1)}(\cdot)$  is the order  $n$  spherical Hankel function [10, 11]. In the spherical coordinate system the fundamental solution has an expansion, called the addition theorem for the fundamental solution, valid for  $|\mathbf{x}| < |\mathbf{y}|$  [9, p.62]

$$\Phi(\mathbf{x}, \mathbf{y}) = ik \sum_{m,n} j_n(k|\mathbf{x}|) h_n^{(1)}(k|\mathbf{y}|) Y_n^m(\hat{\mathbf{x}}) \overline{Y_n^m(\hat{\mathbf{y}})}, \quad (26)$$

from which we glean for  $|x| \leq S < |y|$

$$\beta_{n;s}^m = (-i)^{n-1} k [\mathcal{J}_n(S)]^{1/2} h_n^{(1)}(k|y|) \overline{Y_n^m(\hat{y})}. \quad (27)$$

An arbitrary field can be modelled by a distribution (superposition) of point sources. An arbitrary field which is source-free in  $\mathbb{B}_S^3$  can be modelled by the superposition of point sources located in  $\mathbb{R}^3 \setminus \mathbb{B}_S^3$ .

### F. Orthogonal Projections

In the time-frequency case in the FOBE context, two projection operators play a key role: time-limiting and band-limiting [3]. For spatial systems the analogous operators are more subtle. We define these analogous operators. Later, we will give a fuller justification and generalization.

Consider a function  $w \in L^2(\mathbb{B}_S^3)$ . In general such a function will not correspond to a valid wave-field, that is, it will violate (1). Valid solutions to (1) in  $\mathbb{B}_S^3$  belong to  $\mathcal{G}_S$  which is a linear subspace of  $L^2(\mathbb{B}_S^3)$ . Hence we can define an orthogonal projection operator to map to this subspace.

**DEFINITION 1 (WAVE-FIELD – HELMHOLTZ PROJECTOR).** The Helmholtz Projector  $G_S: L^2(\mathbb{B}_S^3) \rightarrow \mathcal{G}_S$  is defined by

$$w \mapsto u = G_S w \quad (28a)$$

$$G_S w = \sum_{n,m} \langle w, \varphi_{n;s}^m \rangle_{\mathbb{B}_S^3} \varphi_{n;s}^m, \quad (28b)$$

which projects  $w \in L^2(\mathbb{B}_S^3)$  to a solution  $u = G_S w$  to the homogeneous Helmholtz Equation in  $\mathbb{B}_S^3$ .

Since every  $u \in \mathcal{G}_S$  is a fixed point under  $G_S$ , that is  $G_S u = u$ , then  $G_S$  is idempotent,  $G_S^2 = G_S$ . By the Hilbert Space Projection Theorem,  $G_S w$  is the best  $L^2$ -approximation in  $\mathcal{G}_S$  to  $w \in L^2(\mathbb{B}_S^3)$ .

The Helmholtz Projector can be utilized in interesting ways. Consider the situation where a wave-field is measured in the presence of independent gaussian noise. Consequently, one does not have a valid wave-field, in the sense defined above, due to the noise. In this case the Helmholtz Projector can be used to generate the best valid wave-field matching the measurements, in a minimum mean square error sense.

Although not immediately apparent the Helmholtz Projector,  $G_S$ , has an equivalent interpretation but expressed in terms of the source-field  $f(x)$  domain. It is a truncation that removes sources from the ball  $\mathbb{B}_S^3$ :

**DEFINITION 2 (SOURCE-FIELD TRUNCATION – “THE MUMMY” PROJECTOR).** The Mummy projector  $\tilde{G}_S: L^2(\mathbb{R}^3) \rightarrow L^2(\mathbb{R}^3 \setminus \mathbb{B}_S^3)$  is defined by

$$(\tilde{G}_S f)(x) = \begin{cases} f(x) & \text{when } x \in \mathbb{R}^3 \setminus \mathbb{B}_S^3, \\ 0 & \text{otherwise.} \end{cases} \quad (29)$$

As such this projection is parameterized by the radius  $S$ . Clearly this operator is idempotent  $\tilde{G}_S^2 = \tilde{G}_S$  and self-adjoint (as all orthogonal projections must be).

We seek yet another projection to fully furnish the analogy with the time-frequency case. The complementary domain is

the wave-field and the natural projection is the direct truncation of the wave-field (not necessarily one from a homogeneous Helmholtz Equation). That is, we have a projection,  $H_R$ , that takes a function on  $\mathbb{R}^3$  and sets it to zero outside  $\mathbb{B}_R^3$ .

**DEFINITION 3 (WAVE-FIELD TRUNCATION).** The truncation operator  $H_R: L^2(\mathbb{R}^3) \rightarrow L^2(\mathbb{B}_R^3)$  is defined by

$$(H_R w)(x) = \begin{cases} w(x) & \text{when } x \in \mathbb{B}_R^3, \\ 0 & \text{otherwise.} \end{cases} \quad (30)$$

As such this projection is parameterized by the radius  $R$ . Later we will show why these projections are crucial.

### III. FINITE DIMENSIONAL APPROXIMATIONS

As a necessary precursor to finding the essential dimensionality of wavefields in space, we need to develop the appropriate but standard notion of a finite dimensional approximation to what is an infinite dimensional object. Clearly not all finite dimensional approximations are equal and the most important sense in which we choose to distinguish them will be the minimization the error (between the approximation and the original function), using the appropriate norm, given that the dimension is fixed. The essential dimensionality means that beyond some well-defined threshold in dimension, the best finite dimensional approximation achieves a high relative precision *independent of the function to be approximated*.

We now wish to consider the effectiveness of replacing the generic expansion (10) by a finite number of terms in the form

$$u(x) \approx \sum_{p \in P} \gamma_p \varphi_p(x) \quad (31)$$

where  $P$  is a finite index set whose cardinality,  $|P| < \infty$ , is the dimension of the desired representation,  $\gamma_p$  are coefficients to be determined, and  $\{\varphi_p\}_{p \in P}$  is a finite, incomplete, orthonormal set.

Examination of (31) reveals that there are three components which impact on obtaining the best approximation: (i) optimal determination of the orthonormal set  $\{\varphi_p(x)\}_{p \in P}$ , (ii) optimal choice of index set,  $P$ , with a well defined ordering, and (iii) optimal determination of the coefficients for an explicit  $u(x)$  (having chosen the orthonormal set and the index set). Clearly the metric to determine the degree of approximation should be the induced norm (5). The third consideration leads to a standard problem with known solution which is  $\gamma_p = \langle u, \varphi_p \rangle_{\mathbb{B}_S^3}$ , the Fourier Coefficients [7]. Bringing these facts together leads to the consideration of the  $|P|$ -dimension approximation that minimizes

$$\|u(x) - \sum_{p \in P} \langle u, \varphi_p \rangle_{\mathbb{B}_S^3} \varphi_p(x)\|_{\mathbb{B}_S^3} \quad (32)$$

through the choice of  $\{\varphi_p(x)\}_{p \in P}$ .

Determination of the best approximation (32) of a given dimension to represent a wavefield  $u(x)$  in a region  $\mathbb{B}_R^3$ , is non-trivial and resolved in the theoretical developments of the next few sections.

#### IV. SPATIAL CONCENTRATION OF WAVE-FIELDS

##### A. Problem Statement

Now we pose the following problem which is intimately related to the problem of finding the best finite dimensional representation of a wave-field in a ball but it is also of independent interest.

**SPATIAL CONCENTRATION PROBLEM 1.** *Determine wave-field  $u \in \mathcal{G}_S$  — the closed subspace in  $L^2(\mathbb{B}_S^3)$  satisfying  $\Delta u + k^2 u = 0$  in  $\mathbb{B}_S^3$  — which has the greatest concentration of energy in  $\mathbb{B}_R^3$  where  $R < S$ . That is, find the  $u \in \mathcal{G}_S$  which achieves*

$$\sup_{\substack{u \neq 0 \\ u \in \mathcal{G}_S}} \frac{\int_{\mathbb{B}_R^3} |u(\mathbf{x})|^2 dv(\mathbf{x})}{\int_{\mathbb{B}_S^3} |u(\mathbf{x})|^2 dv(\mathbf{x})} \equiv \sup_{\substack{u \neq 0 \\ u \in \mathcal{G}_S}} \frac{\int_{\mathbb{B}_R^3} |u(\mathbf{x})|^2 dv(\mathbf{x})}{\|u\|_{\mathbb{B}_S^3}^2} \quad (33a)$$

or equivalently

$$\sup_{\substack{\|u\|_{\mathbb{B}_S^3}^2 = 1 \\ u \in \mathcal{G}_S}} \int_{\mathbb{B}_R^3} |u(\mathbf{x})|^2 dv(\mathbf{x}). \quad (33b)$$

In plain English, we are want to determine which member(s) of the class of wave-fields generated by sources outside radius  $S$ , of unit energy in the region  $|\mathbf{x}| \leq S$ , have the maximum concentration of energy in the inner ball  $|\mathbf{x}| \leq R$  where  $R < S$ . This is a nontrivial problem as there is no constraint on where the sources of the wave-field may be positioned within  $\mathbb{R}^3 \setminus \mathbb{B}_S^3$ .

We can use the two projections from section II-F to recast Spatial Concentration Problem 1 in a more abstract and compact form:

**SPATIAL CONCENTRATION PROBLEM 2.** *Determine the solution(s) and value of:*

$$\sup_{w \neq 0} \frac{\|H_R G_S w\|_{\mathbb{B}_S^3}^2}{\|G_S w\|_{\mathbb{B}_S^3}^2} \equiv \sup_{\substack{u \neq 0 \\ u \in \mathcal{G}_S}} \frac{\|H_R u\|_{\mathbb{B}_S^3}^2}{\|u\|_{\mathbb{B}_S^3}^2} \quad (34)$$

We will have more to say of the generalization of this formulation in section VII.

##### B. Spatial Concentration Problem Solution

We now solve Spatial Concentration Problem 1. Since  $u$  in (33) is in  $\mathcal{G}_S$  then we can use representation (21) to determine, *without approximation*, the total energy in the inner  $R$ -radius ball

$$\|u\|_{\mathbb{B}_R^3}^2 \triangleq \int_{\mathbb{B}_R^3} |u(\mathbf{x})|^2 dv(\mathbf{x}) \quad (35a)$$

$$= \int_{\mathbb{B}_R^3} \left| \sum_{m,n} \beta_{n;s}^m \varphi_{n;s}^m(\mathbf{x}) \right|^2 dv(\mathbf{x}) \quad (35b)$$

$$= \sum_{m,n} \sum_{p,q} \beta_{n;s}^m \overline{\beta_{q;s}^p} \int_{\mathbb{B}_R^3} \varphi_{n;s}^m(\mathbf{x}) \overline{\varphi_{q;s}^p(\mathbf{x})} dv(\mathbf{x}) \quad (35c)$$

$$= \sum_{m,n} |\beta_{n;s}^m|^2 \frac{\int_0^R [j_n(kr)]^2 r^2 dr}{\mathcal{J}_n(S)} \quad (35d)$$

$$= \sum_{m,n} |\beta_{n;s}^m|^2 \frac{\mathcal{J}_n(R)}{\mathcal{J}_n(S)} \quad (35e)$$

where the  $\beta_{n;s}^m$  are given by (22) and further satisfy  $\|\beta\|_{\ell^2}^2 = 1$  by the imposed unit energy condition  $\|u\|_{\mathbb{B}_S^3}^2 = 1$  in (33b) using (12). In the calculation above we have used the poly-orthogonality property, (16) and (17), of (13), and the orthonormality of the spherical harmonics (14) and (15).

Note that in principle we could have used any complete orthonormal sequence to represent  $u \in \mathcal{G}_S$  but in our calculation the choice of  $\{\varphi_{n;s}^m\}_{m,n}$ , from (21), shows it is optimal in the sense of explicitly finding solutions to (33) as will be revealed shortly. That is, we can solve (33) because we happen to exploit the preferred orthonormal representation; or we could say that we are guided by physical intuition or insight to obtain the preferred orthonormal representation as was done in [6]. This clearly points to the need for a more advanced deductive principle to determine solutions to the Spatial Concentration Problem 1 or its generalizations particularly when we have less symmetries in the formulation. This generalization and theory is the subject of Section VII. In the mean time we will borrow terminology from the more general material.

Define the positive real “eigenvalues”, which have an energy ratio interpretation,

$$\lambda_n(R, S) \triangleq \frac{\mathcal{J}_n(R)}{\mathcal{J}_n(S)} \equiv \frac{\int_0^R [j_n(kr)]^2 r^2 dr}{\int_0^S [j_n(kr)]^2 r^2 dr} \quad (36)$$

for  $n = 0, 1, 2, \dots$ , whereupon we observe maximizing (36) concerns only optimizing the choice of  $n$  as follows

$$\tilde{n} = \arg \max_n \lambda_n(R, S) \quad (37)$$

with no constraint on  $m$  other than  $|m| \leq \tilde{n}$  (with all choices being equal). Hence the Spatial Concentration Problem 1 has a continuum of solutions defined for the  $\tilde{n}$  in (37) by

$$\tilde{u}(\mathbf{x}) \triangleq \sum_{m=-\tilde{n}}^{\tilde{n}} \beta_{\tilde{n};s}^m \varphi_{\tilde{n};s}^m(\mathbf{x}), \text{ where } \sum_{m=-\tilde{n}}^{\tilde{n}} |\beta_{\tilde{n};s}^m|^2 = 1 \quad (38)$$

which is a  $2\tilde{n}+1$  dimensional unit sphere, a subset of the infinite dimensional unit sphere  $\|\beta\|_{\ell^2}^2 = 1$ . The non-uniqueness is expected because of the symmetries of the sphere.

From properties of the spherical Bessel functions on the intervals  $[0, R]$  and  $[0, S]$  where  $R < S$ , one can show [10, 11]

$$\lambda_0(R, S) > \lambda_1(R, S) > \lambda_2(R, S) > \dots > 0 \quad (39)$$

and, therefore, the complete solution to (37) is  $\tilde{n} = 0$ , with  $m = 0$ . Hence the most spatially concentrated wave-field, up to a phase factor, is given by

$$\varphi_{0;s}^0(\mathbf{x}) = \frac{j_0(k|\mathbf{x}|)}{\sqrt{4\pi}[\mathcal{J}_0(S)]^{1/2}} \propto \frac{\sin k|\mathbf{x}|}{k|\mathbf{x}|} \quad (40)$$

which has energy  $\lambda_0$  within  $\mathbb{B}_R^3$  compared with unit energy within  $\mathbb{B}_S^3$ . This solution has the character of an isotropic field which may be generated by uniformly distributed sources in all directions ( $4\pi$  steradians) with perfect phase alignment (ideal constructive interference at the origin).

### C. Optimum Eigenfunction Expansion

The ordering of the eigenvalues in (39) induces an ordering of the orthonormal functions. Given any function  $u \in \mathcal{G}_S$  then we may write

$$u = \langle u, \varphi_{0;s}^0 \rangle \varphi_{0;s}^0 + u' \quad (41)$$

where  $u' \in \mathcal{G}_S^{(1)}$  is orthogonal to  $\varphi_{0;s}^0$ , and we have the direct sum  $\mathcal{G}_S = \text{span}\{\varphi_{0;s}^0\} \oplus \mathcal{G}_S^{(1)}$ . Then replacing  $\mathcal{G}_S$  in the Spatial Concentration Problem 1 with  $\mathcal{G}_S^{(1)}$  leads to a solution associated with  $\lambda_1$ , the next largest eigenvalue after  $\lambda_1$ , that can be expressed as a continuum of functions parametrized by points on a 3 dimensional sphere

$$u(x) \triangleq \sum_{m=-1}^1 \beta_{1;s}^m \varphi_{1;s}^m(x), \text{ where } \sum_{m=-1}^1 |\beta_{1;s}^m|^2 = 1 \quad (42)$$

which has energy  $\lambda_1$  within  $\mathbb{B}_R^3$  compared with unit energy within  $\mathbb{B}_S^3$ . And so on working through the eigenvalues in order from largest to smallest.

We note that it is preferable to take into consideration the multiplicity of the eignvalues when selecting the size of any finite dimensional approximation. In the above, it is natural to choose the dimensions as 1, 4, 9, 16, ... This corresponds to truncating according to  $n \leq N$  and using the full available range for  $m$  ( $|m| \leq n$  for all  $n \leq N$ ). Hence we have determined the best  $\sum_{m,n \leq N} 1 = (N+1)^2$  dimensional approximation in the induced norm, in the form of (31) where  $|P| = (N+1)^2$ :

**THEOREM 1.** *The best  $(N+1)^2$  dimensional approximation of solutions to  $\Delta u + k^2 u = 0$  in  $\mathbb{B}_S^3$  is given*

$$u_N(x) \triangleq \sum_{m,n \leq N} \beta_{n;s}^m \varphi_{n;s}^m(x), \quad (43)$$

where the orthonormal sequence is given by (18). That is, truncation of the normalized modes, (18), in their natural order give the optimal approximations of degree  $(N+1)^2$  for  $\mathbb{B}_S^3$ .

Here  $N$  subscript denotes the truncation depth in  $n$  and not the dimension which is  $(N+1)^2$ .

### V. ESSENTIAL DIMENSIONALITY OF WAVE-FIELDS

We have seen how to construct a finite dimensional approximation to a wave-field, with unit energy in  $\mathbb{B}_S^3$ , that best captures the energy of that wavefield in an inner ball  $\mathbb{B}_R^3$ . Now we show how the accuracy of the best approximation is affected as we vary the dimension by changing  $N$ . We will find that all fields can be well-approximated once  $N$  is chosen large enough and this choice is a function only of the radii  $S$  and  $R$ . That is, there is an essential dimensionality that depends only on the geometry of the region of interest (determined by radius  $R$ ) and the geometry of the sources which generate the wave-field (sources located outside radius  $S$ ). This property was derived in [5,6] in a more restricted context (primarily requiring farfield sources), using different tools and only in the form of bounding the essential dimensionality with uncertainty regarding the tightness.

We begin with  $u_N(x)$ , the optimal approximation given by (43). Using the poly-orthogonality, (16), in analogy with the arguments used in (35), we have

$$\|u - u_N\|_{\mathbb{B}_R^3}^2 \triangleq \int_{\mathbb{B}_R^3} |u(x) - u_N(x)|^2 dv(x) \quad (44a)$$

$$= \int_{\mathbb{B}_R^3} \left| \sum_{m,n > N} \beta_{n;s}^m \varphi_{n;s}^m(x) \right|^2 dv(x) \quad (44b)$$

$$= \sum_{m,n > N} |\beta_{n;s}^m|^2 \frac{\mathcal{J}_n(R)}{\mathcal{J}_n(S)} \quad (44c)$$

$$= \sum_{n > N} \frac{\mathcal{J}_n(R)}{\mathcal{J}_n(S)} \sum_{m=-n}^n |\beta_{n;s}^m|^2. \quad (44d)$$

Since  $\|u\|_{\mathbb{B}_S^3} = 1$  we have  $\|\beta_S\|^2 = 1$  and hence

$$\sum_{n=0}^{\infty} \sum_{m=-n}^n |\beta_{n;s}^m|^2 = 1 \text{ implies } \sum_{m=-n}^n |\beta_{n;s}^m|^2 \leq 1. \quad (45)$$

So we end up with the simple but effective bound:

$$\|u - u_N\|_{\mathbb{B}_R^3}^2 \leq \sum_{n > N} \frac{\mathcal{J}_n(R)}{\mathcal{J}_n(S)} \equiv \sum_{n > N} \lambda_n(R, S) \quad (46a)$$

$$= \sum_{n > N} \frac{\int_0^R [j_n(kr)]^2 r^2 dr}{\int_0^S [j_n(kr)]^2 r^2 dr}. \quad (46b)$$

Hence the bound on the approximation is a function only of  $R$ ,  $S$  and  $N$ , independent of  $u(x)$ . That there is a well-defined and marked threshold as we increase  $N$  is a property that we can infer from the spherical Bessel functions following the bounding arguments given in [5, 6].

### VI. OTHER RESULTS IN SPATIAL CONCENTRATION

#### A. Two Dimensional / Height Invariant Wave-Fields

Results for the two dimensional problem *cannot* be straightforwardly deduced or inferred from the three dimensional case so far considered. However, it is true that a similar general approach can be followed. In this section we point out salient features of the two dimensional case and present results without proof.

There are two interpretations of the two dimensional case: 1) beginning with the two dimensional Helmholtz equation we can develop a theory in the two independent spatial variables and make no reference to a third independent spatial variable, this is mathematically valid but strictly disconnected from physical reality; and 2) the two dimensional case corresponds to a linear subspace of the three dimensional case where the sources and fields are invariant along one spatial dimension, typically taken in the  $z$  direction, the "height", as we assume here. This latter case is physically more relevant and corresponds to a form of cylindrical symmetry where a fundamental solution is interpreted as an infinite line source in the  $z$  direction. Such a line source generates cylindrical waves. However, one can argue such lines sources are only a mathematical idealization. For this reason we express the results using the first interpretation

so that the analogs of  $\mathbb{B}_S^3$  and  $\mathbb{B}_R^3$  are disks  $\mathbb{B}_S^2$  and  $\mathbb{B}_R^2$ , respectively.

The analog of Spatial Concentration Problem 1 has the most spatially concentrated field, up to a constant complex factor, given by

$$J_0(k|\mathbf{x}|) \quad (47)$$

which is the Bessel function of order zero. This two dimensional wave-field contains

$$\frac{\int_0^R [J_0(kr)]^2 r dr}{\int_0^S [J_0(kr)]^2 r dr} \quad (48)$$

of the energy in the disk  $\mathbb{B}_R^2$  given unity energy in  $\mathbb{B}_S^2$ . No other wave-field in  $\mathbb{B}_R^2$  contains more energy (given unit energy in  $\mathbb{B}_S^2$ ). This conclusion, like the three dimensional case, holds irrespective of the values of  $R$  and  $S$ .

The orthonormal sequence which yields the best finite dimensional approximation is the normalization of modes

$$i^n J_n(k|\mathbf{x}|) e^{in\phi(\hat{\mathbf{x}})}, \quad n = 0, \pm 1, \pm 2, \dots \quad (49)$$

where  $|\mathbf{x}|$  is the radius and  $\phi(\hat{\mathbf{x}})$  is the angle in polar coordinates. The dimensionality of truncating the series to  $n < |N|$  is  $2N + 1$ . The essential dimensionality occurs when  $N = e\pi R/\lambda$ . That is the essential dimensionality is given by  $2e\pi R/\lambda + 1$ . Additional details on this case can be found in [5, 6] but in a narrower context.

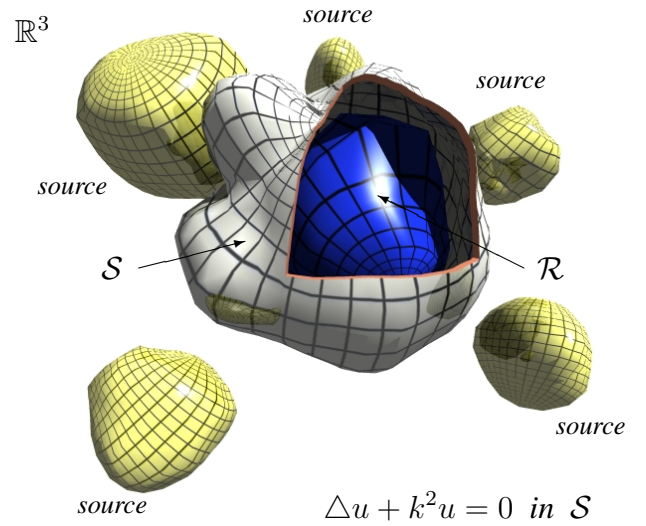
This essential dimensionality therefore, holds for the height invariant wave-field case with respect to the infinite length cylinders of radii  $S$  and  $R$ . That this essential dimensionality is very much less than the case of the finite spheres of radii  $S$  and  $R$  is a manifestation of the strong constraint placed on the wave-field to be height invariant.

### B. One Dimensional / Planar Wave-Fields

The essential dimensionality of the one dimensional Helmholtz equation is two since for a given wavenumber  $k$ , we can only have a one wave traveling left to right and one wave traveling right to left. This represents a degenerate case as the region of interest on the real number is irrelevant (provided it is not of measure zero). In  $\mathbb{R}^3$  the one dimensional Helmholtz equation arises from the consideration of two planar wave fronts whose directions of propagation are opposite.

## VII. GENERAL THEORY OF SPATIAL CONCENTRATION AND DIMENSION

The preceding results for balls which have spherical symmetry represent a special case of a more general result which we now present based on the theory of positive (and self-adjoint) operators in Hilbert Space. In the theory of time-frequency concentration and essential dimensionality the domain and range of the Fourier Transform are the real line  $\mathbb{R}$  and the regions under investigation are either simple closed intervals or a set of disjoint closed intervals. In the wave-field case the domain and range are in general  $\mathbb{R}^3$  (or, as we have seen, perhaps the lower dimensional Euclidean space  $\mathbb{R}^2$  or the degenerate case of  $\mathbb{R}$ ).



**Fig. 2:** Source-field and Wave-field Geometry: The region of interest is the inner blob  $\mathcal{R}$ . Any source for the wave-field, symbolized by the irregular outer blobs, are outside an outer blob  $\mathcal{S}$ , shown with cutaway. That is, the source-field is non-zero only in  $\mathbb{R}^3 \setminus \mathcal{S}$  and  $\mathcal{R} \subset \mathcal{S}$ .

Regions in  $\mathbb{R}^3$  are much richer in structure and the domains become much more tightly coupled into any problem formulation. Such an increase in complexity when more than one independent variable is involved has led to the study of boundary value problems and integral equations [12, 13]. We will use such theory only to the extent of establishing the most important results and leave the obvious extensions and applications to other publications.

We can regard a region in  $\mathbb{R}^3$  as a blob. It is convenient to think that the blob is topologically equivalent to a ball but this is not essential. To apply our results only requires sufficient smoothness conditions on the interior and boundary of the region without any necessity that the blob be simply connected. In the following, the respective blobs  $\mathcal{S}, \mathcal{R} \subset \mathbb{R}^3$  generalize the respective balls  $\mathbb{B}_S^3, \mathbb{B}_R^3 \subset \mathbb{R}^3$  with  $\mathcal{R} \subset \mathcal{S}$ . Because balls capture the essence of blobs we can move rapidly through the concepts associated with blobs without undue elaboration and explanation:

Define the separable Hilbert Space

$$\mathcal{G}_S \triangleq \{u \in L^2(\mathcal{S}) : \Delta u + k^2 u = 0\}. \quad (50)$$

which has inner product

$$\langle f, g \rangle_S \triangleq \int_{\mathcal{S}} f(\mathbf{x}) \overline{g(\mathbf{x})} dv(\mathbf{x}), \quad (51)$$

with induced norm

$$\|f\|_S^2 \triangleq \int_{\mathcal{S}} |f(\mathbf{x})|^2 dv(\mathbf{x}). \quad (52)$$

We now pose the Spatial Concentration Problem:

**SPATIAL CONCENTRATION PROBLEM 3.** Determine wave-field  $u \in \mathcal{G}_S$  — the closed subspace in  $L^2(\mathcal{S})$  satisfying  $\Delta u + k^2 u = 0$  in  $\mathcal{S}$  — which has the greatest concentration



of energy in  $\mathcal{R}$  where  $\mathcal{R} \subset \mathcal{S}$ . That is, find the  $u \in \mathcal{G}_S$  which achieves

$$\sup_{\substack{\|u\|_S=1 \\ u \in \mathcal{G}_S}} \int_{\mathcal{R}} |u(\mathbf{x})|^2 dv(\mathbf{x}). \quad (53)$$

This can be reformulated using operators  $G_S$  and  $H_{\mathcal{R}}$  which are the straightforward generalizations of the source-field and wave-field truncation operators given in section II-F,  $G_S$  and  $H_{\mathcal{R}}$ , respectively:

**SPATIAL CONCENTRATION PROBLEM 4.** *Determine the solution(s) and value of:*

$$\sup_{w \neq 0} \frac{\|H_{\mathcal{R}} G_S w\|_S^2}{\|G_S w\|_S^2} \equiv \sup_{\substack{u \neq 0 \\ u \in \mathcal{G}_S}} \frac{\|H_{\mathcal{R}} u\|_S^2}{\|u\|_S^2}. \quad (54)$$

A closely related problem to (54) is

$$\sup_{w \neq 0} \frac{\|H_{\mathcal{R}} G_S w\|_S^2}{\|w\|_S^2} = \sup_{\|w\|=1} \|H_{\mathcal{R}} G_S w\|_S^2 \quad (55a)$$

$$\triangleq \|H_{\mathcal{R}} G_S\|^2 \quad (55b)$$

which is the square of the norm of the operator  $H_{\mathcal{R}} G_S$  [13].

Now given that  $G_S$  is idempotent, were  $\check{w}$  the supremizer of (55) (if necessary considered as the limit) then, since  $\|G_S w\|_S \leq \|w\|_S$  for all  $w$ , we would find that  $G_S \check{w}$  would also be a supremizer. Hence,  $\check{w}$  needs to be zero in  $\mathbb{R} \setminus \mathcal{S}$ . Then,  $\|G_S \check{w}\|_S^2 = \|\check{w}\|_S^2$ , which shows that a solution to (55) is also a solution to (54) (but not necessarily vice versa). Hence, the solution to these classes of problem reduces to the study of the eigenstructure of the relevant operator equations. The analysis can then proceed as for the time-frequency case. This will be studied in greater depth in a sequel.

### VIII. CONCLUSIONS

The theory developed in the paper determines the maximum concentration of energy that can be placed in an arbitrary shaped region of interest devoid of sources. Associated optimum finite dimensional approximations have been developed. For the case of ball shaped region (where the boundary is a sphere) and where the sources are contained in the complement of a sphere closed form solutions to the problem have been found. These results imply there is a fundamental and calculable limit to the amount of information that may be carried in a region of space since there is an essential limit to the number of orthonormal basis functions that can be used.

A strong analogy was shown to exist with the classical work done on the time-frequency concentration which is tied with the fundamental limits imposed by the uncertainty principle of the Fourier Transform. In the spatial case, the complementary domains are the source-field (distribution of sources which may be nearfield, farfield, point or distributed) and the wave-field it induces. These have the analogies of time and frequency in the case of the Fourier Transform. Further, the analogy of the Fourier Transform is the Helmholtz Equation which is the partial differential equation corresponding to the time independent wave equation.

The theory, which is still being developed underpins the limits placed on communication systems which purport to exploit space. We finish with one example. In MIMO there has been considerable fuss made over the linear increase in capacity of MIMO with numbers of antennas. This is an artifact of imposing independence which is valid if the antennas can be widely separated in space. In the opposite extreme, our theory makes it clear that the capacity is not related to the number of antennas but the region occupied by the antennas. The saturation of performance improvement with number of antennas will come when the essential dimensionality of the region is reached. In the presence of finite precision or noise there will be a grizzly fundamental limit to the information content that a spatial region can bear.

### REFERENCES

- [1] D. Slepian and H. O. Pollack, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - I," *Bell System Tech. J.*, vol. 40, pp. 43-63, Jan. 1961.
- [2] H. J. Landau and H. O. Pollack, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - II," *Bell System Tech. J.*, vol. 40, pp. 65-84, Jan. 1961.
- [3] H. J. Landau and H. O. Pollack, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - III: the dimension of the space of essentially time- and band-limited signals," *Bell System Tech. J.*, vol. 41, pp. 1295-1336, July 1962.
- [4] K. Selig, "Uncertainty principles revisited," *Electron. Trans. Num. Anal.*, 2002, (to appear).
- [5] R. A. Kennedy, T. D. Abhayapala, and H. M. Jones, "Bounds on the Spatial Richness of Multipath," in *Proc. 3rd Australian Communications Theory Workshop, AusCTW'2002*, February 2002, pp. 76-80.
- [6] H. M. Jones, R. A. Kennedy, and T. D. Abhayapala, "On Dimensionality of Multipath Fields: Spatial Extent and Richness," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, ICASSP'2002*, Orlando, Florida, May 2002, vol. 3, pp. 2837-2840.
- [7] F. Riesz and B. Sz. Nagy, *Functional Analysis*, Dover Publications Inc., New York, NY, second edition, 1990.
- [8] L. Debnath and P. Mikusiński, *Introduction to Hilbert Spaces with Applications*, Academic Press, San Diego, CA, 1999.
- [9] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, Berlin, Germany, second edition, 1998.
- [10] N. W. McLachlan, *Bessel Functions for Engineers*, Oxford University Press, London, second edition, 1961.
- [11] C. A. Coulson and A. Jeffrey, *Waves: A Mathematical Approach to the Common Types of Wave Motion*, Longman, London, second edition, 1977.
- [12] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Wiley Interscience, New York, NY, 1953.
- [13] I. Stakgold, *Green's Functions and Boundary Value Problems*, John Wiley & Sons, Inc., New York, NY, second edition, 1998.