Spatial channels for communicating with waves between volumes

David A. B. Miller

Ginzton Laboratory, 450 Via Palou, Stanford University, Stanford, California 94305-4085

Received August 4, 1998

I show that there is an exact, complete method for finding the orthogonal spatial channels, or communications modes, between two arbitrary volumes, and the associated connection strengths, for the case of scalar waves. I also show that the sum of the squared connection strengths is given exactly by a simple volume integral. The method is illustrated by a calculation for a particular extreme pair of volumes, and the communications modes are interpreted physically as the modes of a double phase-conjugate resonator. © 1998 Optical Society of America

OCIS codes: 350.7420, 260.1960, 110.2990, 190.5040.

In optics, the number of resolvable spots, or spatial degrees of freedom, in communicating with waves between two plane-parallel surfaces is relatively well understood; it agrees with a simple diffraction limit to the spot size (see, e.g., Ref. 1). In addition to the intuitive diffraction picture,[1] for rectangular or circular surfaces there is a more rigorous eigenfunction picture[2,3] that also gives the strength of the connection between the two surfaces.

There are, however, many situations in which these approaches break down. The diffraction approach gives no useful answer if either of the volumes is small compared with a wavelength. Both approaches are based on diffraction theory approximations that are not valid in the near field or at large angles. They give no information whatsoever about the effects of depth or thickness and so cannot analyze communication between volumes. Finally, the diffraction approximations used are known to be formally weak (they use only one effective source in a finite aperture rather than the formally required two sources over a closed surface[4]) and are therefore of dubious value for use in drawing fundamental conclusions.

In many situations these approaches, including, for example, any near-field situations (e.g., near-field microscopy, very fine-line lithography) and situations that intrinsically involve volume (e.g., volume holography, some optical interconnects, information storage, and, possibly, modes in photonic bandgap crystals or other volume scatterers), are of limited use. Here it is shown,[5] at least for scalar waves, that it is possible to define, without approximations, the communications channels or modes between arbitrary volumes, to determine their connection strengths, and to derive a general sum rule for the connection strengths between these volumes.

Consider two volumes, a transmitting volume V_T that contains some spatially distributed source $\psi(\mathbf{r}_T)$ and a receiving volume V_R with a wave $\phi(\mathbf{r}_R)$ that is generated by $\psi(\mathbf{r}_T)$. Here we consider only scalar, monochromatic sources and waves, with angular frequency $\omega = kc$, where c is the wave-propagation velocity, which obey the scalar Helmholtz equation

$$\nabla^2\phi(\mathbf{r}) + k^2\phi(\mathbf{r}) = -\psi(\mathbf{r}). \tag{1}$$

Because we consider only outgoing waves from volume V_T, the resultant Green's function that describes the wave at $\mathbf{r}$ from a point source at position $\mathbf{r}_T$ is

$$G(\mathbf{r}, \mathbf{r}_T) = \exp(-ik|\mathbf{r} - \mathbf{r}_T|)/4\pi|\mathbf{r} - \mathbf{r}_T|; \tag{2}$$

so, formally, the total resultant wave is

$$\phi(\mathbf{r}) = \int_{V_T} G(\mathbf{r}, \mathbf{r}_T)\psi(\mathbf{r}_T)\,\mathrm{d}^3\mathbf{r}_T. \tag{3}$$

We want to know the pairs of functions for $\psi(\mathbf{r}_T)$ and $\phi(\mathbf{r}_R)$ that give the best possible couplings between volumes V_T and V_R and define independent (and, ideally, orthogonal) communications channels, or communications modes. For our mathematical approach we define complete orthonormal basis sets of functions, $a_{T1}(\mathbf{r}_T)$, $a_{T2}(\mathbf{r}_T)$, $a_{T3}(\mathbf{r}_T), \ldots$ in V_T and $a_{R1}(\mathbf{r}_R)$, $a_{R2}(\mathbf{r}_R)$, $a_{R3}(\mathbf{r}_R), \ldots$ in V_R. Choosing $\psi(\mathbf{r}_T) = a_{Ti}(\mathbf{r}_T)$ leads to a wave in V_R of the form $\phi(\mathbf{r}_R) = \sum_j g_{ji} a_{Rj}(\mathbf{r}_R)$, with

$$g_{ji} = \int_{V_R}\int_{V_T} a_{Rj}^*(\mathbf{r}_R) G(\mathbf{r}_R, \mathbf{r}_T) a_{Ti}(\mathbf{r}_T)\,\mathrm{d}^3\mathbf{r}_T\,\mathrm{d}^3\mathbf{r}_R. \tag{4}$$

The g_{ji} are coupling coefficients between the transmission from mode i in volume V_T and reception by mode j in volume V_R, giving the amplitude of the wave function $a_{Rj}(\mathbf{r}_R)$ in V_R that results from source $a_{Ti}(\mathbf{r}_T)$ in V_T. We can also show that, if we set up a weak receiving source in V_R of amplitude $\propto a_{Rj}(\mathbf{r}_R)$, the real and the imaginary parts of g_{ji} are essentially the power-coupling coefficients between source function $a_{Ti}(\mathbf{r}_T)$ and receiving source $a_{Rj}(\mathbf{r}_R)$ for the two possible orthogonal phases.

The best choice of (normalized) function $\psi(\mathbf{r}_T)$ would be the one that gave rise to the maximum value of $|g|^2 = \int_{V_R} |\phi(\mathbf{r}_R)|^2\,\mathrm{d}^3\mathbf{r}_R$; i.e., substituting from Eq. (3), we wish to choose $\psi(\mathbf{r}_T)$ to maximize

$$|g|^2 = \int_{V_T} \psi^*(\mathbf{r}_T{}') \int_{V_T} K(\mathbf{r}_T{}', \mathbf{r}_T)\psi(\mathbf{r}_T)\,\mathrm{d}^3\mathbf{r}_T{}'\,\mathrm{d}^3\mathbf{r}_T, \tag{5}$$

$$K(\mathbf{r}_T{}', \mathbf{r}_T) = \int_{V_R} G^*(\mathbf{r}_R, \mathbf{r}_T{}')\, G(\mathbf{r}_R, \mathbf{r}_T) \mathrm{d}^3\mathbf{r}_R\,. \tag{6}$$

[It is straightforward to show that g is the coupling coefficient in the sense of Eq. (4) between normalized functions of the forms $\psi(\mathbf{r}_T)$ and $\phi(\mathbf{r}_R)$.]

Note that $K(\mathbf{r}_T{}', \mathbf{r}_T)$ is Hermitian, i.e., $K(\mathbf{r}_T{}', \mathbf{r}_T) = K^*(\mathbf{r}_T, \mathbf{r}_T{}')$, and, at least with V_R and V_T as completely separate and finite volumes, it is a continuous, finite function. These are sufficient conditions to make the eigenfunctions associated with $K(\mathbf{r}_T{}', \mathbf{r}_T)$ be a complete set in V_T and have real eigenvalues.[6] It is also a standard result that the Hermitian form [Eq. (5)] is maximized when $\psi(\mathbf{r}_T)$ is chosen as the eigenfunction of $K(\mathbf{r}_T{}', \mathbf{r}_T)$ with the largest eigenvalue,[6,7] i.e., the function $\psi(\mathbf{r}_T)$ that solves the integral equation

$$|g|^2 \psi(\mathbf{r}_T{}') = \int_{V_T} K(\mathbf{r}_T{}', \mathbf{r}_T) \psi(\mathbf{r}_T)\, \mathrm{d}^3\mathbf{r}_T \tag{7}$$

with the largest value of $|g|^2$. Indeed, the set of orthonormal functions with the largest successive values of $|g|^2$ is the set of normalized eigenfunctions $\psi_1(\mathbf{r}_T)$, $\psi_2(\mathbf{r}_T)$, $\psi_3(\mathbf{r}_T)$, …, arranged in descending order of their eigenvalues, $|g_1|^2$, $|g_2|^2$, $|g_3|^2$, … .

Now, associated with each $\psi_n(\mathbf{r}_T)$ there is also a normalized function $\phi_n(\mathbf{r}_R)$ for which

$$g_n \phi_n(\mathbf{r}_R) = \int_{V_T} G(\mathbf{r}_R, \mathbf{r}_T) \psi_n(\mathbf{r}_T)\, \mathrm{d}^3\mathbf{r}_T\,. \tag{8}$$

From Eq. (5), interchanging the integration order, and substituting from Eq. (8), we have

$$g_n^* \psi_n(\mathbf{r}_T) = \int_{V_R} G^*(\mathbf{r}_R, \mathbf{r}_T) \phi_n(\mathbf{r}_R)\, \mathrm{d}^3\mathbf{r}_R\,. \tag{9}$$

Using Eq. (9) to substitute for $\psi_n(\mathbf{r}_T)$ in Eq. (8), we find that $\phi_n(\mathbf{r}_R)$ are the solutions of

$$|g_n|^2 \phi_n(\mathbf{r}_R) = \int_{V_R} J(\mathbf{r}_R, \mathbf{r}_R{}') \phi_n(\mathbf{r}_R{}')\, \mathrm{d}^3\mathbf{r}_R{}'\,, \tag{10}$$

$$J(\mathbf{r}_R, \mathbf{r}_R{}') = \int_{V_T} G(\mathbf{r}_R, \mathbf{r}_T) G^*(\mathbf{r}_T, \mathbf{r}_R{}')\, \mathrm{d}^3\mathbf{r}_T\,, \tag{11}$$

which, similarly, have a complete set of eigenfunctions with real eigenvalues.

Hence we have shown that, for any pair of (separate) volumes, there is a set of communications modes that are pairs of functions $[\psi_n(\mathbf{r}_T), \phi_n(\mathbf{r}_R)]$. These are the best coupled functions between the two volumes, and each set, $\psi_n(\mathbf{r}_T)$ and $\phi_n(\mathbf{r}_R)$, is complete and hence is capable of describing any function within its volume.[8]

It is also true that, by taking the complex conjugate of Eq. (9), if we had a source of the form $\phi_n^*(\mathbf{r}_R)$ in volume V_R it would generate a wave of the form $\psi_n^*(\mathbf{r}_R)$ in volume V_T. Suppose that we start with source $\psi_n(\mathbf{r}_T)$ in V_T, which results in the wave $\phi(\mathbf{r}_R) = g_n \phi_n(\mathbf{r}_R)$ in V_R, and now have some phase-conjugating process in V_R that generates a new source of amplitude $A\phi^*(\mathbf{r}_R) = A g_n^* \phi_n^*(\mathbf{r}_R)$. Then the wave generated in volume V_T is $\phi'(\mathbf{r}_T) = A|g_n|^2 \psi_n^*(\mathbf{r}_T)$. If we then similarly have a mechanism in V_T that generates a new source of amplitude $B\phi'^*(\mathbf{r}_T)$, the resultant source is $BA^*|g_n|^2 \psi_n(\mathbf{r}_T)$, which is simply a constant times the original source. Hence this system constitutes a resonator, and the communications modes are essentially the modes of this double phase-conjugate resonator.

The second important result of this approach is to derive a sum rule for the connection strengths. We expand $G(\mathbf{r}_R, \mathbf{r}_T)$ bilinearly in the complete sets $a_{Ti}^*(\mathbf{r}_T)$ and $a_{Rj}(\mathbf{r}_R)$. The expansion coefficients are simply g_{ji} from Eq. (4), which yield

$$G(\mathbf{r}_R, \mathbf{r}_T) = \sum_{i,j} g_{ji} a_{Rj}(\mathbf{r}_R) a_{Ti}^*(\mathbf{r}_T)\,. \tag{12}$$

Using Eq. (12) and integrating over both volumes V_T and V_R, we obtain

$$\int_{V_R} \int_{V_T} |G(\mathbf{r}_R, \mathbf{r}_T)|^2\, \mathrm{d}^3\mathbf{r}_T\, \mathrm{d}^3\mathbf{r}_R = \sum_{i,j} |g_{ji}|^2\,; \tag{13}$$

hence, using Eq. (2), we obtain the sum rule

$$\gamma_{RT} \equiv \sum_{i,j} |g_{ji}|^2 = \frac{1}{(4\pi)^2} \int_{V_R} \int_{V_T} \frac{1}{|\mathbf{r}_R - \mathbf{r}_T|^2}\, \mathrm{d}^3\mathbf{r}_T\, \mathrm{d}^3\mathbf{r}_R \tag{14}$$

(which becomes a diagonal sum if we use the communications modes as the basis sets). This sum rule, Eq. (14), has the remarkable property that it depends only on the shape and relative position of the two volumes. Loosely, it tells us that, though there are in general an infinite number of communications modes, there is a limit on the product of the number of channels and their usefulness.

With the sum rule and the communications mode solutions we can derive the known results[2] for plane, rectangular surfaces far apart, by formally analyzing the case of thin cuboidal volumes, although this derivation is not shown here. Such analysis also shows that the thickness does not increase the number of useful channels for communication between such distant volumes. (It is important to emphasize that this result is for the communications channels for information going in and out of the volume and not for the number of degrees of freedom for storage of information in the volume, which is quite a different calculation.)

Figure 1 shows an extreme case with two very thin (0.1-wavelength) volumes at right angles and only one wavelength apart. For calculation I chose a finite Fourier basis in each volume (and took the solutions to be constant in the thin directions), obtaining a finite matrix of coefficients; such finite basis approaches give an underestimate of the coupling strengths.[6] The first mode [Fig. 1(b)] takes at least ~86% of the available communications strength (i.e., $|g_1|^2 = 0.86\ \gamma_{RT}$), and the second mode [Fig. 1(c)] takes at least ~11% (i.e., $|g_1|^2 = 0.11\ \gamma_{RT}$). The sum rule tells us that we have accounted for all the strongly connected modes, there being only ~3% of γ_{RT} left. The generated waves are relatively obviously orthogonal to one another; the source functions are orthogonal also, although it is

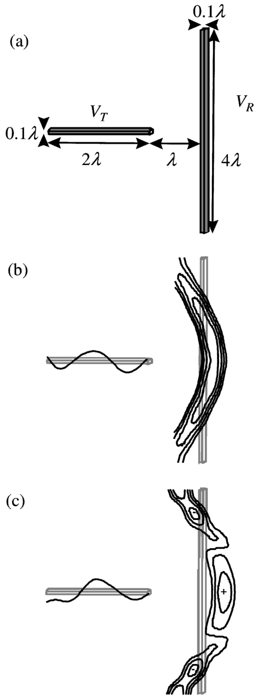

Fig. 1. Illustrations (a) of two thin volumes considered in this example, (b) the strongest communications mode, and (c) the second communications mode. For the transmitting volume the real part of the wave amplitude along the length of the volume is shown for a particular arbitrary phase. For the receiving volume, the real part of the wave is shown in a contour plot that illustrates approximately half a period of the wave and with a horizontal scale such that 2π of phase is the same size as one wavelength on the diagram. With this choice of scale the curvature of the phase fronts corresponds approximately to the actual curvature of the propagating wave. λ is the wavelength.

necessary to look at the entire complex function to see this clearly. Note that, in this extreme case, there is a second communications mode associated with the depth of the transmitting volume.

In this example we have at least five different extreme conditions, any of which would have violated the previous surface models; here (i) the volumes are smaller than a wavelength in at least one lateral dimension, (ii) the volumes are closer together than the depth of one volume, (iii) one of the volumes is presented edge-on to the other, (iv) one volume subtends a large solid angle to the other, and (v) the volumes are only a wavelength apart.

In conclusion, it has been shown that there is an exact, complete way of analyzing the communications modes for scalar waves between two arbitrary volumes, yielding explicit formulas for calculating those modes. The sum rule that we derived shows that, although there may be an infinite number of such modes, only a finite number can be strongly connected. This overall approach may help in the analysis of wave problems for which the usual diffraction approaches are invalid for any of a number of reasons.

References

1. D. Gabor, in *Progress in Optics*, E. Wolf, ed. (North-Holland, Amsterdam, 1961), Vol. 1, pp. 109–153.
2. G. Toraldo di Francia, J. Opt. Soc. Am. **59,** 799 (1969).
3. General results have also been derived for arbitary shapes of apertures in the eigenfunction picture in the study of propagation through atmospheric turbulence; see, e.g., J. H. Shapiro, in *Laser Beam Propagation in the Atmosphere*, J. W. Strohbehm, ed. (Springer-Verlag, Berlin, 1978), pp. 171–222.
4. It is possible to use one source type rigorously over a closed surface for the case of phase fronts; see D. A. B. Miller, Opt. Lett. **16,** 1370 (1991).
5. Portions of this research were presented in D. A. B. Miller, Proc. SPIE **3490,** 111–114 (1998).
6. See, e.g., D. Porter and D. S. G. Stirling, *Integral Equations* (Cambridge U. Press, Cambridge, 1990), for a discussion of these issues.
7. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Wiley Classics ed. (Wiley, New York, 1989).
8. This process of finding the communications modes is analogous to singular value decomposition of the matrix g_{ji}; to prove the existence and completeness of the sets of eigenfunctions, however, requires results from functional analysis and hence an integral representation of the problem.