Devising genuine spatial correlation functions

F. Gori* **and M. Santarsiero**

Università Roma Tre, Dipartimento di Fisica, and CNISM, Via della Vasca Navale 84, I-00146 Rome, Italy
**Corresponding author: gori@uniroma3.it*

Received September 19, 2007; revised November 19, 2007; accepted November 19, 2007;
posted November 21, 2007 (Doc. ID 87642); published December 10, 2007

The choice of the mathematical form of spatial correlation functions for optical fields is restricted by the constraint of nonnegative definiteness. We discuss a sufficient condition for ensuring the satisfaction of such a constraint. © 2007 Optical Society of America

OCIS codes: 030.1640, 260.5430.

In coherence theory there are many instances in which we wish to figure out the mathematical form of the spatial correlation functions of a partially coherent and partially polarized field. In doing so, we have to make sure that such functions correspond to nonnegative definite kernels [1]. Except for certain cases, such as those pertaining to Schell-model sources [1,2], this may be a difficult task. Therefore, any sufficient condition that ensures satisfaction of the nonnegative definiteness requirement is welcome. In this Letter, we present a sufficient condition of this type, derived from the theory of reproducing kernel Hilbert spaces.

The spatial coherence properties of an optical field are generally described by means of either the mutual intensity [1,3–5] or the cross-spectral density (CSD) [1,5]. Both of these functions have to be nonnegative definite kernels. To recall the meaning of this phrase, let $W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)$ be the CSD across a (primary or secondary) planar source at points $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$. For brevity, we do not explicitly show the dependence of W_0 on the temporal frequency. Now, define the quantity

$$Q(f)=\int\int W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)f^*(\boldsymbol{\rho}_1)f(\boldsymbol{\rho}_2)\mathrm{d}^2\rho_1\mathrm{d}^2\rho_2, \tag{1}$$

where $f(\boldsymbol{\rho})$ is a well-behaving function. Then, nonnegative definiteness means that $Q(f)$ has to be greater than or equal to zero for any choice of f. It should also be recalled that, for any nonnegative definite kernel, the following inequality holds:

$$|W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)|^2\le W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_1)W_0(\boldsymbol{\rho}_2,\boldsymbol{\rho}_2). \tag{2}$$

In optics, $W_0(\boldsymbol{\rho},\boldsymbol{\rho})$ coincides with the spectral density $S_0(\boldsymbol{\rho})$ across the source plane [1,5].

While it may be simple to check that inequality $Q(f)\ge 0$ is satisfied for some selected functions f, to guarantee that this happens for any choice of f is not that easy. In principle, we could check whether a kernel is nonnegative by performing a modal analysis [1,5]. In fact, all the eigenvalues of a nonnegative kernel are nonnegative. Unfortunately, this requires the (generally infinitely many) solutions of a homogeneous Fredholm integral equation of the second kind to be found in closed form. Except for a limited number of cases, this cannot be done.

Sometimes the violation of the nonnegativeness condition can be indirectly revealed by absurd physical results. For example, a certain form of the function W_0 might predict a negative value for the spectral density at some point in space. This is enough to exclude such a function from consideration. As we shall see in a moment, more subtle cases may occur.

We work out an example, limiting ourselves, for simplicity, to sources for which the CSD depends on one transverse coordinate only, say ξ, for each point. Accordingly, the CSD is denoted $W_0(\xi_1,\xi_2)$.

Let us recall that there exist sources for which the CSD takes on the factorized form

$$W_0(\xi_1,\xi_2)=F\left(\frac{\xi_1+\xi_2}{2}\right)G(\xi_1-\xi_2), \tag{3}$$

F and G being suitable functions. Gaussian Schell-model sources [1,5] constitute the most celebrated case for which the CSD can be exactly written in this way [6]. The present expression is also used for describing the CSD of quasi-homogeneous sources [1,5], but in this case it is meant to be an approximation of the true CSD. Here, we are interested in sources for which the above expression is an exact representation. Some conditions under which this occurs were established in [7]. Now, let us assume F and G to be a Lorentzian and a Gaussian function, respectively. More explicitly, with reference to a typical variable s, we let

$$F(s)=\frac{S_M}{1+\beta s^2},\quad G(s)=e^{-\alpha s^2}, \tag{4}$$

where S_M, α, and β are positive constants. On inserting from Eqs. (4) into Eq. (3), we find the following expression for the CSD:

$$W_0(\xi_1,\xi_2)=\frac{S_M e^{-\alpha(\xi_1-\xi_2)^2}}{1+\beta(\xi_1+\xi_2)^2/4}. \tag{5}$$

We now wonder, can this be the exact expression of a CSD, at least under a suitable choice for the parameters α and β? Let us write the spectral density and the spectral degree of spatial coherence [1,5] across the source plane. They are given by

$$S_0(\xi)=W_0(\xi,\xi)=\frac{S_M}{1+\beta\xi^2}, \tag{6}$$

$$\mu_0(\xi_1,\xi_2) = \frac{W_0(\xi_1,\xi_2)}{\sqrt{S_0(\xi_1)S_0(\xi_2)}} = \frac{\sqrt{(1+\beta\xi_1^2)(1+\beta\xi_2^2)}}{1+\beta(\xi_1+\xi_2)^2/4} e^{-\alpha(\xi_1-\xi_2)^2}. \quad (7)$$

It is possible to show that to have $\mu_0(\xi_1,\xi_2)$ not exceeding one for any choice of ξ_1 and ξ_2 it is required that $\alpha/\beta \geq 1/4$.

Let us now consider propagation in the paraxial regime. Following the lines presented in [8], it can be seen that the spectral density across a plane at a distance z from the source can be evaluated through the formula

$$S_z(x) = \frac{1}{\lambda z}\int F(\sigma)\tilde{G}\left(\frac{x-\sigma}{\lambda z}\right)\mathrm{d}\sigma, \quad (8)$$

where λ is the wavelength and the tilde denotes the Fourier transform. Since F and $\tilde{G}$ are nonnegative, the predicted spectral density is nonnegative throughout the paraxial domain. Things seem to go smoothly. As a further check, however, let us evaluate the spectral degree of spatial coherence in the far zone, say $\mu_z^{(\infty)}(x_1,x_2)$. There the CSD, say $W_z^{(\infty)}(x_1,x_2)$, is given by [1,5]

$$W_z^{(\infty)}(x_1,x_2) = \frac{1}{\lambda z}\tilde{F}\left(\frac{x_2-x_1}{\lambda z}\right)\tilde{G}\left(\frac{x_1+x_2}{2\lambda z}\right). \quad (9)$$

For the present case, taking Eqs. (4) into account, we find

$$\mu_z^{(\infty)}(x_1,x_2) = \exp\left[-\frac{2\pi}{\sqrt{\beta}}\left|\frac{x_2-x_1}{\lambda z}\right| + \frac{\pi^2}{4\alpha}\left(\frac{x_2-x_1}{\lambda z}\right)^2\right]. \quad (10)$$

Here, we realize that something went wrong. As a matter of fact, for $|x_2-x_1| > 8\alpha\lambda z/(\pi\sqrt{\beta})$, the spectral degree of spatial coherence would become greater than one, further growing without limit on increasing $|x_2-x_1|$. This nonsensical result shows that, in spite of its plausible structure, the mathematical form (5) for the CSD has to be rejected. It may be noted that a spectral degree of spatial coherence greater than one would imply that inequality (2) is violated. This is just another way of asserting that Eq. (5) cannot represent a genuine CSD [9].

We are thus led back to the basic question: how can we conceive correct CSDs? In other words, are there sufficient conditions ensuring that a kernel $W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)$ built according to some recipe is nonnegative definite?

The same question is encountered in the theory of reproducing kernel Hilbert spaces [10–14]. Such a theory has been used in several subjects, such as probability and statistics, signal processing, numerical analysis, and machine learning, but, to our knowledge, has not yet been exploited in coherence theory. Some nonnegative definiteness criteria can be found in books and papers relating to reproducing kernel Hilbert spaces. Among them, the most useful for our purposes seems to be the following. A kernel W_0 is nonnegative definite if it can be written as a superposition integral of the form

$$W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = \int p(\boldsymbol{v})H^*(\boldsymbol{\rho}_1,\boldsymbol{v})H(\boldsymbol{\rho}_2,\boldsymbol{v})\mathrm{d}^2v, \quad (11)$$

where H is an arbitrary kernel and p is a nonnegative function [15]. The function H is assumed to be well behaving, whereas p can also consist of a set of δ functions (with positive coefficients). Then, on using Eq. (11) in Eq. (1), the quantity Q takes on the form

$$Q(f) = \int p(\boldsymbol{v})\left|\int H(\boldsymbol{\rho},\boldsymbol{v})f(\boldsymbol{\rho})\mathrm{d}^2\rho\right|^2\mathrm{d}^2v, \quad (12)$$

which clearly shows Q to be nonnegative for any choice of f, provided that $p(\boldsymbol{v})$ is everywhere nonnegative.

A representation of the form of Eq. (11) was proposed, on physical grounds, in [16], by generalizing a procedure first used in [17] in connection with Gaussian Schell-model sources. More explicitly, there exists a class of partially coherent sources that can be synthesized through the incoherent superposition of mutually shifted replicas of a basic coherent field $V(\boldsymbol{v})$ with weights $w(\boldsymbol{v})$. The CSD then assumes the form

$$W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = \int w(\boldsymbol{v})V^*(\boldsymbol{\rho}_1-\boldsymbol{v})V(\boldsymbol{\rho}_2-\boldsymbol{v})\mathrm{d}^2v, \quad (13)$$

where $w(\boldsymbol{v})$ is supposed to be a nonnegative function (or a set of positive δ functions). Clearly, this is of the form of Eq. (11) when we let $H(\boldsymbol{\rho},\boldsymbol{v}) = V(\boldsymbol{\rho}-\boldsymbol{v})$ and $p(\boldsymbol{v}) = w(\boldsymbol{v})$. It should be noted that, in a recent paper, Vahimaa and Turunen [18] illustrated the usefulness of representation (13) for studying propagation processes.

It is worthwhile to consider the case in which V has a Gaussian form, namely, $V(\boldsymbol{\rho}) = \exp(-2\alpha\rho^2)$, with α a positive constant. On inserting such an expression into Eq. (13), the following form is obtained for the CSD:

$$W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = F\left(\frac{\boldsymbol{\rho}_1+\boldsymbol{\rho}_2}{2}\right)G(\boldsymbol{\rho}_1-\boldsymbol{\rho}_2), \quad (14)$$

with

$$F(\boldsymbol{s}) = \int p(\boldsymbol{v})\exp[-4\alpha(\boldsymbol{s}-\boldsymbol{v})^2]\mathrm{d}^2v \quad (15)$$

and $G(\boldsymbol{s}) = \exp(-\alpha s^2)$. We see that Eq. (14) has the same form as Eq. (3). The term depending on the difference of the coordinates is Gaussian. The F function is the convolution between a Gaussian and an arbitrary positive weight function p. For instance, the use of a Gaussian function for the weight function would lead to the familiar case of a Gaussian Schell-model source, but an infinite class of sources would be obtained on varying p. As a remark, we note that no weight function can be found such that its convolution with a Gaussian gives rise to a Lorentzian function, as can be easily verified, so that the CSD of our first example [Eq. (4)] does not satisfy the sufficient condition (13) (written for a one-dimensional case).

Coming back to the general superposition scheme, Eq. (11), we observe that an intuitive view of it can be gained as follows. Suppose that the source plane $z=0$ actually coincides with the output plane of a linear system endowed with a coherent impulse response $H(\boldsymbol{\rho},\boldsymbol{v})$. Furthermore, imagine that at the input of such system a field distribution characterized by a CSD $W_i(\boldsymbol{v}_1,\boldsymbol{v}_2)$ is present. Then, the CSD $W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)$ is easily seen to be

$$W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = \int\int W_i(\boldsymbol{v}_1,\boldsymbol{v}_2)H^*(\boldsymbol{\rho}_1,\boldsymbol{v}_1)H(\boldsymbol{\rho}_2,\boldsymbol{v}_2)\mathrm{d}^2v_1\mathrm{d}^2v_2. \tag{16}$$

Now, if the input field distribution is spatially incoherent, i.e., if its CSD has the form $W_i(\boldsymbol{v}_1,\boldsymbol{v}_2) = p(\boldsymbol{v}_1)\delta(\boldsymbol{v}_1-\boldsymbol{v}_2)$, we see at once that Eq. (16) reduces to Eq. (11). We should not stick to this view too literally, though. Indeed, from the mathematical point of view, $H(\boldsymbol{\rho},\boldsymbol{v})$ represents the kernel of an arbitrary linear transformation and could be different from physically available coherent impulse responses.

The use of the sufficient condition (11) allows us to investigate structures of correlation functions that appear to be rather new with respect to the forms usually encountered in coherence theory.

As a class of examples, let us refer to a kernel H that depends on $\boldsymbol{v}$ through a Fourier-like exponential, i.e.,

$$H(\boldsymbol{\rho},\boldsymbol{v}) = \tau(\boldsymbol{\rho})\exp[-2\pi i\boldsymbol{v}\cdot\boldsymbol{g}(\boldsymbol{\rho})], \tag{17}$$

where $\tau(\boldsymbol{\rho})$ is a (possibly complex) profile function, whereas $\boldsymbol{g}(\boldsymbol{\rho})$ is a real vector function. Then, using any arbitrary nonnegative, Fourier transformable function $p(\boldsymbol{v})$, Eq. (11) leads to

$$W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = \tau^*(\boldsymbol{\rho}_1)\tau(\boldsymbol{\rho}_2)\tilde{p}[\boldsymbol{g}(\boldsymbol{\rho}_1)-\boldsymbol{g}(\boldsymbol{\rho}_2)]. \tag{18}$$

In particular, suppose $\boldsymbol{g}\,(\boldsymbol{\rho})=a\boldsymbol{\rho}$, where a is a real parameter. Then, Eq. (18) becomes

$$W_0(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = \tau^*(\boldsymbol{\rho}_1)\tau(\boldsymbol{\rho}_2)\tilde{p}[a(\boldsymbol{\rho}_1-\boldsymbol{\rho}_2)]. \tag{19}$$

This shows that Eq. (18) includes, as a particular case, Schell-model sources. Note that the parameter a merely plays the role of a scale factor within $\tilde{p}$. However, Eq. (18) gives rise to a much wider set of correlation functions because $\boldsymbol{g}(\boldsymbol{\rho})$ can be chosen at will.

Obviously, countless other classes could be constructed. For our aim, what has been said hitherto is sufficient to suggest how large a variety of genuine CSDs can be explored with the help of sufficient condition (11).

As a further remark, note that p could be made of a finite or denumerably infinite set of δ functions with positive weights p_v. In this case, the variable v plays the role of an index. If we choose p_v as coincident with vth eigenvalue of the source [1], and the kernel with the corresponding mode, this resulting expression of the CSD reduces to the modal expansion of the source.

In conclusion, we have seen that by exploiting a basic result from the theory of reproducing kernel Hilbert spaces we can establish sufficient conditions [see Eq. (11)] for devising *bona fide* spatial correlation functions. This can be of help for exploring new forms of correlation functions and for putting into evidence some unexpected features of them. We limited ourselves to the scalar case, but analogous results can be obtained in the vector case, i.e., when polarization effects have to be taken into account.

References

1. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge U. Press, 1995).
2. P. De Santis, F. Gori, G. Guattari, and C. Palma, J. Opt. Soc. Am. A **3**, 1258 (1986).
3. M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge U. Press, 1999).
4. J. W. Goodman, *Statistical Optics* (Wiley, 1983).
5. E. Wolf, *Introduction to the Theory of Coherence and Polarization of Light* (Cambridge U. Press, 2007).
6. F. Gori, Opt. Commun. **46**, 149 (1983).
7. F. Gori and E. Wolf, Opt. Commun. **61**, 369 (1987).
8. F. Gori, Opt. Lett. **4**, 354 (1979).
9. The authors are indebted to A. T. Friberg for calling this point to their attention.
10. N. Aronszajn, Trans. Am. Math. Soc. **68**, 337 (1950).
11. E. Parzen, Ann. Math. Stat. **32**, 951 (1961).
12. G. Wahba, *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics 59 (Society for Industrial and Applied Mathematics, 1990).
13. S. Saitoh, *Integral Transforms, Reproducing Kernels and Their Applications* (Addison Wesley Longman, 1997).
14. A. Berlinet and C. Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics* (Kluwer Academic, 2003).
15. For any nonnegative definite kernel there exists a unique functional Hilbert space whose functions are reproduced when they are multiplied by the kernel by means of the scalar product holding for such space. This is the origin of the phrase “Reproducing kernel Hilbert space.”
16. F. Gori, Opt. Acta **27**, 1025 (1980).
17. F. Gori and C. Palma, Opt. Commun. **27**, 185 (1978).
18. P. Vahimaa and J. Turunen, Opt. Express **14**, 1376 (2006).