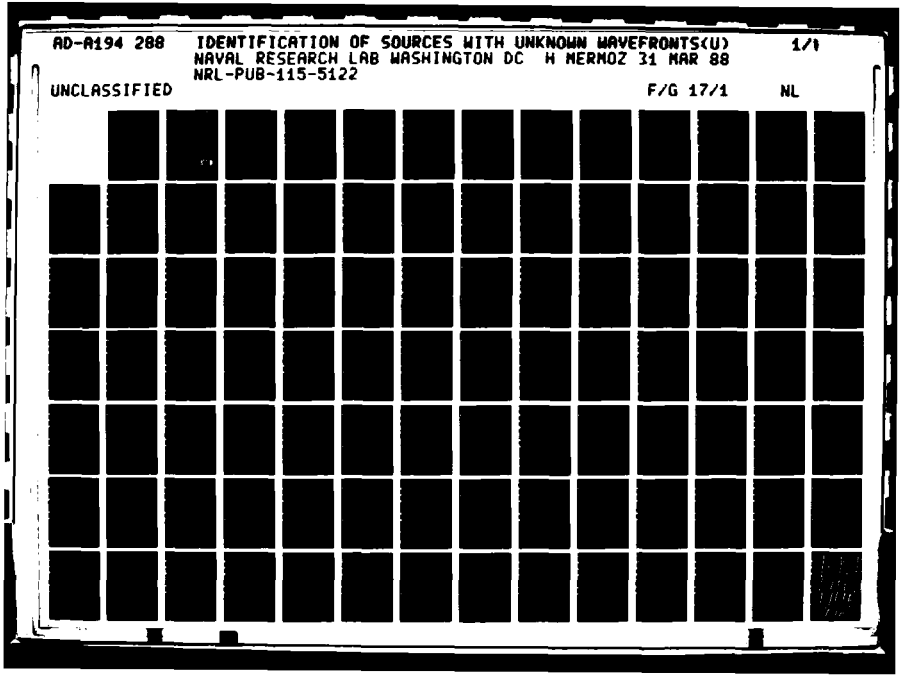
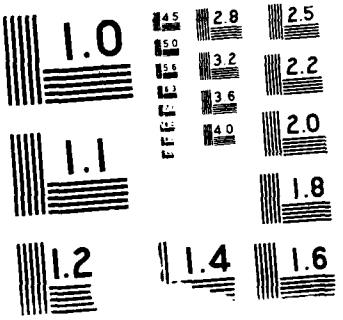

AD-A194 288 IDENTIFICATION OF SOURCES WITH UNKNOWN WAVEFRONTS(U) 1/1
NAVAL RESEARCH LAB WASHINGTON DC H MERMOZ 31 MAR 88
NRL-PUB-115-5122
UNCLASSIFIED F/G 17/1 NL


1.0
1.1
1.2
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0

Naval Research Laboratory

Washington, DC 20375-5000 Publication 115-5122

Identification of Sources with Unknown Wavefronts

HENRI MERMOZ

March 31, 1988

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704 0188

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Ocean Research and Development Activity	6b OFFICE SYMBOL (If applicable) 5120
6c ADDRESS (City, State, and ZIP Code) Stennis Space Center, MS 39529-5004	7b ADDRESS (City, State, and ZIP Code)
8a NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research	8b OFFICE SYMBOL (If applicable)
8c ADDRESS (City, State, and ZIP Code) Arlington, VA 22217-5000	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER 10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 61153N; PROJECT NO: RR032-04-04; TASK NO: ; WORK UNIT ACCESSION NO:

11 TITLE (Include Security Classification)

Identification of Sources with Unknown Wavefronts (U)

12 PERSONAL AUTHOR(S)
Mermoz, Henri

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year Month Day)	15 PAGE COUNT
Final	FROM ___ TO ___	1988 March 31	95

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Directivity; Source Identification; Wavefronts; Adaptive Signal Processing; Cross-Spectral Density Matrix

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

(U) This paper concerns techniques for processing sonar array data in order to obtain information about the spatial distribution of acoustic sources. The goal is to present a method for determining the spectral density and the signal wavefronts of several sources using a minimum of prior assumptions. Consideration of this problem is the next logical step along the evoloutionary path that spatial processing has taken over the last few decades. This path is characterized by a decreasing level of confidence in any kind of prior knowledge about the effects of the propagation medium, which has proven to be more and more elusive and unpredictable as detection ranges and array size have increased.

(Continued)

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL Peter Mignerey	22b TELEPHONE (Include Area Code) 202-767-2230
	22c OFFICE SYMBOL 5122

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

(U) The passive sonar problem is defined and the traditional approach to a solution using matched filters is presented. Maximization of the directivity factor for known signals with static, nonadaptive filters emerges as the prime goal of traditional processing. Care is then taken to properly introduce the cross-spectral density matrix which plays a crucial role in adaptive processing.

(U) A realization of adaptive noise cancellation through the use of signal-free references is achieved with the development of a "Correlo-Filtre". Unfortunately adaptive noise cancellation leads to degradation when a supposedly known wavefront deviates from an expected shape. Robustness and the need for a processor which does not require prior knowledge of the wavefront is then discussed.

(U) A novel method of identifying completely unknown wavefronts is proposed. Using the eigenvalues of the cross-spectral density matrix allows the number of sources to be determined. The wavefront shapes are then shown be scrambled within the eigenvectors by an unknown unitary matrix. Additional information needed to select a specific unitary matrix is introduced by modeling the sources as a coherent sum of "conditional plane waves". While the existence of a solution is posed, no specific solution is obtained because of the large computational effort required.

DD Form 1473, JUN 86

Preface

Dr. Mermoz has worked in the area of acoustic signal processing since the end of World War II. As the first director of the French naval laboratory Groupe Etudes Recherches Détection Sousmarines, he guided the development of French acoustic signal processing from the ashes of war to one of the premier laboratories in existence today. Now retired, this paper represents a desire of Dr. Mermoz to share with us his knowledge of the past, understanding of the present, and vision of the future in passive sonar signal processing. Hence this report is divided into three main chapters (2–4) concerned with each evolutionary stage in our understanding of acoustic signal processing.

In chapter 2 he defines the passive sonar problem and presents the *traditional* approach to a solution using matched filters. Maximization of the directivity factor for known signals with static, nonadaptive filters emerges as the prime goal of traditional processing. He then takes some care to properly introduce the cross-spectral density matrix which plays a crucial role in adaptive processing discussed in the remainder of this report.

In chapter 3 he turns his attention to contemporary methods for the adaptive cancelation of noise through the use of signal-free references. He achieves a realization of noise cancelation with the development of a "Correlo Filtre". Unfortunately adaptive noise cancelation works too well, leading to degradation when a supposedly known wavefront deviates from what is expected. He then proceeds to discuss robustness and the need for a processor which does not require prior knowledge of the wavefront.

Thus we are lead into chapter 4 in which Dr. Mermoz puts forth a novel method of identifying completely unknown wavefronts. Using the eigenvalues of the cross-spectral density matrix allows the number of sources to be determined. The wavefront shapes are then shown be scrambled within the eigenvectors by an unknown unitary matrix. Additional information

needed to select a specific unitary matrix is introduced by modeling the sources as a coherent sum of "conditional plane waves". While the existence of a solution is posed, no specific solution is obtained because of the large computational effort required. This then is his vision of one future area of profitable research in sonar array processing.

Some mention must be made of the nonstandard notation used to differentiate between various types of vectors and matrices. Row vectors accented with a left arrow, e.g. $\vec{\mathbf{R}}$, are used throughout this paper to denote complex filter coefficients. Column vectors accented with a right arrow, e.g. $\vec{\mathbf{S}}$, are used to denote a complex wavefront. Matched filters tuned to a particular wavefront are represented by the complex conjugate transpose of the wavefront, $\vec{\mathbf{S}}^{\dagger}$. Matrices are accented with a double arrow, $\ddot{\mathbf{C}}$.

Cross-spectral density matrices (CSDM) of different types are differentiated using subscripts. Throughout sections 2.1–3.4 $\ddot{\mathbf{C}}$ represents noise alone, while in the remainder of the paper $\ddot{\mathbf{C}}$ represents a transformed noise alone CSDM. The CSDM of perfect sources is represented by $\ddot{\mathbf{C}}_s$, while the raw data matrix is denoted by $\ddot{\mathbf{C}}_r$.

The Naval Research Laboratory contracted Dr. Mermoz to write this report in August 1983. The finished report was delivered to NRL on 20 June 1986. Since that time the paper has been revised on a part-time basis by the editor.

Any technical correspondence concerning this paper should be addressed to the author at: 582 Chemin de la Calade, 83140 SIX-FOURS, France.

Peter Mignerey,
Editor

Acknowledgement

The author expresses his thanks: to Dr. Raymond Fitzgerald of ONR who, while still working for NRL, suggested this contract, did the work necessary to get it granted, and edited chapter 2; to the Office of Naval Research for funding this work; to the NRL authoritites who agreed to monitor it, particularly Mr. Burton Hurdle; most of all, to Dr. Peter Mignerey of NRL who toiled over a manuscript in approximate English and whose fruitful criticisms have greatly improved the presentation.

Contents

List of Figures

Chapter 1

INTRODUCTION

This paper deals with techniques for processing sonar array data in order to obtain information about the spatial distribution of the acoustic sources. The goal is to present a method for identifying sources whose frequency spectrum and wavefront at the array are unknown. Consideration of this problem is the next logical step along the evolutionary path that spatial processing has taken over the last few decades. This path is characterized by a decreasing level of confidence in any kind of prior knowledge about the *effects of the propagation* medium, which has proven to be more and more elusive and unpredictable as detection ranges and array size have increased.

In order to provide a proper introduction for this problem, it is useful to retrace the evolutionary path with a modern point of view so that the new concepts can be progressively introduced. As a result, this introduction is divided into three sections, which correspond roughly to the three major steps in the evolution of spatial signal processing.

The concept of directivity index has dominated (if not choked) spatial processing for decades. In fact, the utility of directivity index as a descriptor requires very stringent conditions to be satisfied: both the signal wavefront and the spatial structure of the noise field must be known. The signal wavefront is usually assumed to be a plane wave, which implies that the propagation medium is homogeneous. The noise field is usually assumed to be "omnidirectional," which again implies a homogeneous medium and, in addition, completely determines the noise correlation as a function of the spatial separation between array sensors. Medium behavior is therefore frozen in the notion of directivity index; the only option left for spatial

processing is to try to improve the signal to noise ratio, which depends on the number of sensors and on their geometry.

Directivity index was a satisfactory descriptor for the high-frequency short-range sonars of an earlier generation. Many senior scientists, however, can remember a time when the sonar equation—in decibels—was allowed an additional term, A, for "anomaly." Because the medium was frozen and allowed for no abnormalities, the concept had to evolve.

Circa 1960 it became obvious that a description of the noise field as omnidirectional was no longer satisfactory. Moreover, it became apparent that the description of the spatial structure as frozen was to be questioned. The new picture of a time-varying spatial structure called for the development of time-varying processors. This led to the introduction of adaptive spatial processing and adaptive beamforming, in which *any thought of using prior knowledge about the noise field is dismissed.* Instead the description of the noise field reflected in the time-dependent correlation matrix changes with time. Here the notion of correlation matrix appears explicitly, whereas it was only implicit in the concept of directivity index. Although the correlation matrix plays an important role in adaptive beamforming, that role remains limited compared to what it will be in the approach presented in chapter 4.

Although the notion of a predetermined, static noise field had been relaxed, the signal wavefront was presumed still to be a plane wave whose direction of propagation was known. The resulting beamformer operated to look simultaneously in all possible directions with preformed beams. In chapter 3 we will generalize the problem slightly by assuming that, although the signal wavefront is known, it is not necessarily a plane wave. "Beamforming" then becomes adaptive processing matched to a given signal wavefront.

Adaptive beamforming was doubtless a significant step forward in that it eliminated all prior assumptions about the structure of the noise field. Its cost lies in the need for a significant amount of real time computing power and the techniques for its implementation are still developing.

It was predictable, and it has already been shown, that adaptive beamforming—even in the generalized form described in chapter 3—would be limited by the only assumption it retained: namely, that the signal wavefront is known. At first glance, any attempt to eliminate this last assumption would seem hopeless. Nevertheless, a deeper analysis of the structure

of the correlation matrix, shows that, *although some prior information is required*, the amount is less than that contained in a complete description of the signal wavefront. Moreover, assumption of a too detailed prior description of the signal wavefront can easily prove to be incompatible with the measured correlation matrix. The temptation always exists, in that case, to blame the incompatibility on inaccuracies in the estimated correlation matrix. That this is not the source of the problem is suggested both by our increasing ability to acquire accurate data and by the development of refined estimation methods. We shall comment only briefly on the copious literature devoted to estimation methods: essentially we shall accept and trust these methods as they exist. Having done so we can consider attempting to cope with the complexity of the medium and its effects in order to determine both the signal wavefront and its spectral density with large arrays. This is still adaptive processing, but not simply beamforming, and it requires a huge amount of real time computing power.

What we are attempting to do, then, is to present a method for the "*identification*" (i.e. determination of the spectral density and the signal wavefronts) of one or more sources using a minimum of prior assumptions. The very notions of signal and noise will be slightly different from what they were in chapters 2 and 3.

We shall not deal with the problem of source localization nor shall we deal with any specific model of the medium. Instead we shall indicate in chapter 4 what can be done in approaching the problem of source localization (i.e. determining the coordinates of the source). In previous papers, most of which are in French, we mixed the problems of identification and localization as is usually done in more superficial treatments. What we shall show here is that localization, and only localization, is inseparable from some sort of description or model of the medium. Separating the two problems will shed clarity on both. While the problem of wavefront identification can be a matter strictly for array processing specialists, the problem of localization follows as an additional study in which those who model the medium should participate.

No specific algorithms will be proposed; we shall merely present the set of equations which must be attacked in order to obtain a practical solution. Given the complexity of the suggested approach, no operational solution is to be expected in the near future. This study should be considered part of a long term approach which may well require new advances in digital

computing power before any real-time implementation becomes possible. Nevertheless, some experimental investigation is possible today and those particular aspects of the theory which require early verification will be indicated. Indeed the kind of "weak" assumptions which remain in chapter 4 are, in fact, the last ramparts in spatial processing. If these or some similar assumptions were totally unacceptable for a given problem, the implication would be that no sensor combination would be more effective than a single sensor. Such a catastrophic case seems quite unlikely, however, because arrays have thus far proven to be useful tools which do provide some gain, although generally not the gain predicted by current theories.

Chapter 2

DIRECTIVITY AND SOURCE IDENTIFICATION

2.1

Much has been written about directivity, yet the concept actually contains two different notions. The more familiar notion of angular resolving power, which is related to beamwidth, has significance for both transmitting and receiving arrays. The scientific literature dealing with the minimization of beamwidth and the reduction of sidelobe levels is extensive. For receiving arrays it is well known that, by using nonlinear combinations of sensor outputs, one can obtain narrower main beams and hence higher resolution than with linear techniques, at least for strong signals. Such nonlinear processing techniques are inefficient and unstable and can lead to erroneous results for weak signals in a noisy environment. The second notion contained in the concept of directivity is that of directivity factor or, in decibels, directivity index, which is a measure of the spatial array gain (i.e. the increase in the signal-to-noise ratio) for signals, weak or strong, in the presence of noise.

Because our intention is to treat the detection and identification of weak signals in background noise, we are concerned here only with the second notion of directivity and with linear combinations of sensors. Thus, we deal with processors in which each sensor output is passed through a linear filter. Each filter can be described in terms either of its impulse response or of its complex frequency response, which form a Fourier pair. The term "combination" will be used in this paper to refer to the sum of the outputs

of these linear filters.

In a nonadaptive beamformer, such as one having classical preformed beams, the filters do not change with time; in adaptive beamformers the filter characteristics are continually adjusted in accordance with the information obtained from estimates of the evolving noise correlation matrix. The permanence of the filters in classical beamforming is a direct consequence of the very rigid assumptions made about the structure of the signal and noise. These assumptions are:

1. The signal is uncorrelated with the noise and is carried by a single plane wave coming from a direction specified by ω which stands for the two angles, bearing and tilt.

2. The noise is omnidirectional (i.e. spherically isotropic). This means that for each frequency the noise field in the vicinity of the array is composed of a superposition of uncorrelated plane waves which can be described as carrying equal power from all directions or as being uniformly distributed over all solid angles $d\omega$. Implicit in both assumptions is the notion that the medium is isotropic.

2.2

For the signal and noise fields described above, the behavior of the array is completely described by its plane wave response function. For a given sensor combination the plane wave response is the complex function $D(\omega, f)$ which gives the amplitude and phase at the beamformer output, when the input is a unit amplitude plane wave having frequency f and coming from the direction ω. The output phase is measured relative to one of the sensors or relative to some reference point in the array.

For a nonrandom signal coming from the direction ω_0, the system response at frequency f is $s(f)D(\omega_0, f)$, where $s(f)$ is the complex amplitude of the Fourier component at frequency f measured at the signal origin. For this signal the power spectral density at the output is $|s(f)|^2 |D(\omega_0, f)|^2$ where the vertical bars denote the modulus of a complex quantity, so that $|D|^2 = DD^*$ with the star denoting complex conjugation.

If the signal coming from the ω_0 direction is stochastic with power spectral density $d(f)$, then the power spectral density at the beamformer output

is $d(f)|D(\omega_0, f)|^2$. The fraction of the isotropic noise in the solid angle $\delta\omega$ contributes a term $b(f)|D(\omega, f)|^2\delta\omega$ to the output spectral density, where $b(f)$, the power spectral density of the noise per unit solid angle, is independent of ω. Because the noise contributions from different solid angles are assumed to be uncorrelated, the total noise power

$$b(f)\int_{\Omega}|D(\omega, f)|^2\,\delta\omega$$

is just the sum of the contributions from the solid angles Ω. For the isotropic noise field under consideration $\Omega = 4\pi$.

2.3

For the stochastic signal coming from direction ω_0, the signal-to-noise ratio at the output of the beamformer can be expressed as a function of frequency as

$$\frac{d(f)|D(\omega_0, f)|^2}{b(f)\int_{\Omega}|D(\omega, f)|^2\,\delta\omega}$$

which is a real, positive quantity proportional to the traditional directivity factor.

Design of a spatial processor then requires maximizing

$$\frac{|D(\omega_0, f)|^2}{\int_{\Omega}|D(\omega, f)|^2\,\delta\omega}$$

by the choice of sensor geometry, of sensor combination, or of both. Often the sensor geometry is fixed and only the choice of sensor combination is available. In that case the functional form of D is known and the solution becomes straightforward. We shall develop the solution in a manner which will make it possible to introduce concepts of importance in later sections of this paper.

The scalar function $D(\omega, f)$ describes two stages of processing:

1. The first stage of processing is performed by the array itself independent of the sensor combination. For a *unit amplitude plane wave* having frequency f and coming from the direction ω, the Fourier transformed sensor outputs provide a set of N complex numbers,

where N is the number of sensors. This set of numbers is the translation through the array of the incoming plane wave and can be considered to be a vector in an N-dimensional complex space. For any given array geometry this vector is easily constructed. Let $h_n(\omega, f)$ be the Fourier-transformed output of the n^{th} sensor evaluated at frequency f and let $H(\omega, f)$ be the vector whose N components are the h_n.

2. The second stage of processing is performed by the combination that is, by summing the outputs of the N linear filters. These filters are described by their frequency responses, $r_n(f)$, which are the components of a vector $\vec{\mathbf{R}}(f)$.

The beamformer output resulting from these two stages of processing is

$$D(\omega, f) = \sum_{n=1}^{N} h_n(\omega, f) r_n(f). \tag{2.1}$$

Understanding that all quantities are functions of frequency, we shall hence forth suppress the frequency variable in order to simplify the notation.

If $\vec{\mathbf{R}}$ is taken to be a row vector and $\vec{\mathbf{H}}(\omega)$ a column vector, then (2.1) becomes the scalar product

$$D(\omega) = \vec{\mathbf{R}}\vec{\mathbf{H}}(\omega). \tag{2.2}$$

The column vector $\vec{\mathbf{H}}(\omega)$ is a particular case of what we shall call a "source vector" corresponding to a remote source transmitting white noise which arrives at the array from the direction ω. The corresponding normalized vector is $\vec{\mathbf{F}}(\omega) = \vec{\mathbf{H}}(\omega)/|\vec{\mathbf{H}}(\omega)|$ where $|\vec{\mathbf{H}}(\omega)| = |\vec{\mathbf{H}}^{\dagger}(\omega)\vec{\mathbf{H}}(\omega)|^{1/2}$ and the superscript † indicates the conjugate transpose. We shall refer to the vector $\vec{\mathbf{F}}$ as the "wavefront vector," although it is not the physical wavefront (here a plane wave) but its translation through the array of sensors. In the concept of directivity, it is implicitly assumed that this wavefront vector is perfectly known for plane waves arriving at the given array geometry.

The filter response r_n must be chosen to maximize the ratio

$$\rho(\omega_0) = \frac{|D(\omega_0)|^2}{\int_{\Omega} |D(\omega)|^2 \, \delta\omega}. \tag{2.3}$$

Optimization of $\rho(\omega_0)$ usually leads to a main beam pointed in the ω_0 direction. If, however, additional constraints are placed on the solution,

such as equality of side-lobe levels or a predetermined rate of decrease in side lobe level, then the maximum response of the array may not be in the direction of the source. In other words, optimization of the directivity factor does not necessarily meet the requirements suggested by a desire to improve angular resolution. Still, optimization of the directivity factor does provide the maximum signal-to-noise ratio under the given assumptions. This is one illustration of the previously mentioned duality in the concept of directivity.

2.4

With $D(\omega)$ given by (2.1), the integral in the denominator of $\rho(\omega_0)$ can be written as

$$\int_\Omega D(\omega)D^*(\omega)\,\delta\omega = \int_\Omega \left[\sum_{n=1}^{N}\sum_{m=1}^{N} h_n(\omega)h_m^*(\omega)r_n r_m^*\right]\delta\omega, \tag{2.4}$$

which can be rewritten as

$$\int_\Omega |D(\omega)|^2\,\delta\omega = \sum_{n=1}^{N}\sum_{m=1}^{N}\left[r_n r_m^* \int_\Omega h_n(\omega)h_m^*(\omega)\,\delta\omega\right]. \tag{2.5}$$

The quantity $h_n(\omega)h_m^*(\omega)$, which is the element in the n^{th} row and m^{th} column of the elementary matrix $\vec{\mathbf{H}}(\omega)\vec{\mathbf{H}}^\dagger(\omega)$, is the cross-spectral density of the field coming from the directions ω. This can be seen by recalling the general rule:

> If a source with spectral density σ passes through two filters h_n and h_m in parallel, then the cross-spectral density of the two outputs is $q_{nm} = \sigma h_n h_m^*$ (see appendix A).

In the particular case under consideration $\sigma = 1$, because $\vec{\mathbf{H}}$ corresponds to a unit amplitude plane wave. Thus, the cross-spectral density matrix for a point-like source in the far field is a dyad of the form $\dot{\mathbf{H}}(\omega)\dot{\mathbf{H}}^\dagger(\omega)$. In order to understand the approach in the later sections of this paper, it is essential to observe that the rank of this matrix is 1, even though its order is N. Indeed, every vector $\dot{\mathbf{X}}$ that is orthogonal to $\dot{\mathbf{H}}$ in an $(N-1)$-dimensional subspace is transformed by $\dot{\mathbf{H}}\dot{\mathbf{H}}^\dagger$ to a null vector. Obviously,

with $\vec{\mathbf{H}}$ and $\vec{\mathbf{X}}$ as column vectors, we have $\vec{\mathbf{H}}^\dagger\vec{\mathbf{X}} = 0$ which implies that $(\vec{\mathbf{H}}\vec{\mathbf{H}}^\dagger)\vec{\mathbf{X}} = 0$. The only nonzero eigenvalue of $\vec{\mathbf{H}}\vec{\mathbf{H}}^\dagger$ is the scalar $\vec{\mathbf{H}}^\dagger\vec{\mathbf{H}}$ and the corresponding eigenvector is $\vec{\mathbf{H}}$ itself.

2.5

Looking now at the integral on the right hand side of (2.5), we write

$$c_{nm} = \int_\Omega h_n(\omega) H_m^*(\omega)\, \delta\omega. \tag{2.6}$$

This is clearly the $(n,m)^{\text{th}}$ element in the sum of an infinite number of elementary cross-spectral density matrices linked to uncorrelated point-like sources. Therefore, this sum is the $(n,m)^{\text{th}}$ element of the cross-spectral density matrix $\vec{\mathbf{C}}$ for the *normalized* omnidirectional noise. Although each elementary contribution to matrix $\vec{\mathbf{C}}$ is a rank-one matrix, the rank of $\vec{\mathbf{C}}$ is not unity. Neither is it less than the order N, because the rank increases every time an elementary contribution is added, and there is an infinity of such contributions (this will be better illustrated in chapter 4). As a result, the rank and order of $\vec{\mathbf{C}}$ are N and $\vec{\mathbf{C}}$ is invertible (i.e. $\vec{\mathbf{C}}^{-1}$ exists). Something similar would happen if the sensor outputs consisted of independent noise with the same spectral density. The cross-spectral density matrix would again be of rank N, but it would have a simpler form, being proportional to a unit matrix (which is not true for matrix $\vec{\mathbf{C}}$).

It is obvious from (2.6) that $c_{nm} = c_{mn}^*$. The matrix $\vec{\mathbf{C}}$ is Hermitian, $\vec{\mathbf{C}} = \vec{\mathbf{C}}^\dagger$, and as a result has real, positive eigenvalues, as does every cross-spectral density matrix. The same is true for the inverse matrix $\vec{\mathbf{C}}^{-1}$. The right hand side of (2.5) can now be written as

$$\vec{\mathbf{R}}\vec{\mathbf{C}}\vec{\mathbf{R}}^\dagger = \sum_{n=1}^{N}\sum_{m=1}^{N} r_n r_m c_{nm}$$

which is the quadratic form for the row vector $\vec{\mathbf{R}}$ through the matrix $\vec{\mathbf{C}}$. This real, positive scalar is the spectral density of the omnidirectional noise at the output of the combination of sensors.

2.6

Looking now to the numerator of $\rho(\omega_0)$ in (2.3),

$$|D(\omega_0)|^2 = D(\omega_0)D^*(\omega_0)$$

we can use (2.1),

$$D(\omega_0) = \sum_{n=1}^{N} h_n(\omega_0) r_n,$$

to write

$$D(\omega_0)D^*(\omega_0) = \sum_{n=1}^{N}\sum_{m=1}^{N} r_n r_m^* h_n(\omega_0) h_m^*(\omega_0).$$

The quantity $h_n(\omega_0)h_m^*(\omega_0)$ is the element in the n^{th} row and m^{th} column of the matrix $\vec{\mathbf{H}}_0\vec{\mathbf{H}}_0^\dagger$ related to the incoming plane wave signal. The rank of this matrix is unity. The only nonzero eigenvalue is the scalar $\vec{\mathbf{H}}_0^\dagger\vec{\mathbf{H}}_0$ and the associated eigenvector is $\vec{\mathbf{H}}_0$. As we shall see later, $\vec{\mathbf{H}}_0$ is the source vector of a remote source transmitting a white noise from the direction ω_0. We have

$$|D(\omega_0)|^2 = \vec{\mathbf{R}}\vec{\mathbf{H}}_0\vec{\mathbf{H}}_0^\dagger\vec{\mathbf{R}}^\dagger = |\vec{\mathbf{R}}\vec{\mathbf{H}}_0|^2 \tag{2.7}$$

which is simultaneously the quadratic form of the vector $\vec{\mathbf{R}}$ through the matrix $\vec{\mathbf{H}}_0\vec{\mathbf{H}}_0^\dagger$ and the square of the modulus of the scalar product $\vec{\mathbf{R}}\vec{\mathbf{H}}_0$.

Maximizing $\rho(\omega_0)$ in (1.3) then consists of choosing the filters $\vec{\mathbf{R}}$ so as to maximize

$$\rho(\omega_0) = \frac{|\vec{\mathbf{R}}\vec{\mathbf{H}}_0|^2}{\vec{\mathbf{R}}\vec{\mathbf{C}}\vec{\mathbf{R}}^\dagger}. \tag{2.8}$$

This is exactly the same problem that is encountered in determining the spatial matched filter for a nonrandom signal in the presence of noise, whose complex spectral components at the sensor outputs are $\vec{\mathbf{H}}_0$, with a cross-spectral density matrix given by $\vec{\mathbf{C}}$.

The solution is well known and can be obtained in several ways. One of them is briefly outlined here. In order to present it in classical form, let us change notation by replacing the row matrix $\vec{\mathbf{R}}$ with the row matrix $\vec{\mathbf{X}}^\dagger$. The column matrix $\vec{\mathbf{X}}$ then represents the filters r_n through the convention

$x_n = r_n^*$. In the new notation (2.8) becomes:

$$\rho(\omega_0) = \frac{\left|\vec{\mathbf{X}}^\dagger \vec{\mathbf{H}}_0\right|^2}{\vec{\mathbf{X}}^\dagger \vec{\mathbf{C}} \vec{\mathbf{X}}}. \tag{2.9}$$

Matrix $\vec{\mathbf{C}}$, being Hermitian and invertible, can be written in factored form as $\vec{\mathbf{C}} = \vec{\mathbf{M}}^\dagger \vec{\mathbf{M}}$ with $\vec{\mathbf{M}}$ an invertible matrix. Hence:

$$\vec{\mathbf{C}}^{-1} = \vec{\mathbf{M}}^{-1} \left(\vec{\mathbf{M}}^\dagger\right)^{-1} = \vec{\mathbf{M}}^{-1} \left(\vec{\mathbf{M}}^{-1}\right)^\dagger. \tag{2.10}$$

Defining the column vectors $\vec{\mathbf{Y}}$ and $\vec{\mathbf{P}}_0$ by:

$$\begin{aligned} \vec{\mathbf{Y}} &= \vec{\mathbf{M}} \vec{\mathbf{X}}, \\ \vec{\mathbf{P}}_0 &\equiv \left(\vec{\mathbf{M}}^\dagger\right)^{-1} \vec{\mathbf{H}}_0; \end{aligned} \tag{2.11}$$

equation (2.9) becomes:

$$\rho(\omega_0) = \frac{\left|\vec{\mathbf{Y}}^\dagger \vec{\mathbf{P}}_0\right|^2}{\vec{\mathbf{Y}}^\dagger \vec{\mathbf{Y}}}. \tag{2.12}$$

Now the Schwartz inequality states that for any $\vec{\mathbf{Y}}$ and $\vec{\mathbf{P}}_0$:

$$\left|\vec{\mathbf{Y}}^\dagger \vec{\mathbf{P}}_0\right|^2 \leq \left(\vec{\mathbf{Y}}^\dagger \vec{\mathbf{Y}}\right)\left(\vec{\mathbf{P}}_0^\dagger \vec{\mathbf{P}}_0\right). \tag{2.13}$$

Equality occurs when $\vec{\mathbf{Y}}$ and $\vec{\mathbf{P}}_0$ are parallel:

$$\vec{\mathbf{Y}} = \alpha \vec{\mathbf{P}}_0, \tag{2.14}$$

with α an arbitrary frequency dependent complex scalar. From (2.12) and (2.13), the maximization of $\rho(\omega_0)$ is reached through (2.14). This maximum in conjunction with (2.10) and (2.11) allow $\vec{\mathbf{X}}$ and $\vec{\mathbf{R}}$ to be written as:

$$\begin{aligned} \vec{\mathbf{X}} &= \alpha \vec{\mathbf{C}}^{-1} \vec{\mathbf{H}}_0, \\ \vec{\mathbf{R}} &= \alpha^* \vec{\mathbf{H}}_0^\dagger \vec{\mathbf{C}}^{-1}; \end{aligned} \tag{2.15}$$

which is a spatial generalization of the matched filter theorem. The maximinimum value of $\rho(\omega_0)$ is then,

$$\begin{aligned} \rho_{max} &= \vec{\mathbf{P}}_0^\dagger \vec{\mathbf{P}}_0 \\ &= \vec{\mathbf{H}}_0^\dagger \vec{\mathbf{C}}^{-1} \vec{\mathbf{H}}_0. \end{aligned} \tag{2.16}$$

The filters $\dot{\mathbf{R}}$ and $\dot{\mathbf{X}}$ defined by (2.15) are multiplied by an arbitrary filter factor which does not alter ρ_{max} (provided the filter has no null in its frequency response across the useful bandwidth). If some of the filters are noncausal it is possible to make them so by finding a time delay (or a phase filter in series: $\exp[2\pi f\tau]$), which translates all the impulse responses to causality. Moreover, it is also possible to get an *unaltered* signal spectral density on the output of the filter combination. Since the ultimate goal is to analyze the spectral density, an unaltered signal is generally required. One has to choose $|\dot{\mathbf{R}}_0\dot{\mathbf{H}}_0| = 1$, which is equivalent to

$$|D(\omega_0, f)|^2 = 1, \tag{2.17}$$

because of (2.7). Consequently,

$$\begin{aligned} |o| &= \left(\dot{\mathbf{H}}_0^{\dagger}\ddot{\mathbf{C}}^{-1}\dot{\mathbf{H}}_0\right)^{-1} \\ &= 1/\rho_{max}. \end{aligned} \tag{2.18}$$

A little more restrictive condition $D(\omega_0, f) = 1$, for all ρ, generally makes *the more familiar directivity functions* all-pass filters for the signal. Anyway the final expression for $\dot{\mathbf{X}}$ and $\dot{\mathbf{R}}$ after (2.15) and (2.18) is, within an arbitrary phase factor:

$$\begin{aligned} \dot{\mathbf{X}} &= \frac{\ddot{\mathbf{C}}^{-1}\dot{\mathbf{H}}_0}{\dot{\mathbf{H}}_0^{\dagger}\ddot{\mathbf{C}}^{-1}\dot{\mathbf{H}}_0} \\ \dot{\mathbf{R}} &= \frac{\dot{\mathbf{H}}_0^{\dagger}\ddot{\mathbf{C}}^{-1}}{\dot{\mathbf{H}}_0^{\dagger}\ddot{\mathbf{C}}^{-1}\dot{\mathbf{H}}_0}. \end{aligned} \tag{2.19}$$

2.7

In this section we are concerned with directivity and omnidirectional noise. Vector $\dot{\mathbf{H}}$ is the array response to a unit amplitude plane wave (the geometry of the array), while $\ddot{\mathbf{C}}$ denotes a normalized cross-spectral density matrix linked to $\mathbf{H}$ by (2.6). It is noteworthy that all relations from (2.9) to (2.19) are valid for any vector $\dot{\mathbf{H}}_0$ and any cross-spectral density matrix $\ddot{\mathbf{C}}$. For both a nonrandom impulse signal, $\delta(t)$ "carried by the wavefront" $\dot{\mathbf{H}}_0$, and a stochastic white signal with a unit spectral density also carried by $\dot{\mathbf{H}}_0$, the

optimal set of filters is given by (2.19) when the noise spatial structure is represented by the nonnormalized matrix $\ddot{\mathbf{C}}$. It happens that we have also partly treated the processing problem for any spatial signal represented by a given vector in a noise structure represented by a given matrix. This will be used in chapter 3.

Coming back to directivity, we now have an output with the unaltered signal spectral density d mixed with the fraction of the omnidirectional noise which is let through by the combination $\dot{\mathbf{R}}$ or $\dot{\mathbf{X}}$. According to section 2.3, this noise spectral density is:

$$\begin{aligned} b\int_{\Omega}|D(\omega)|^2\,\delta\omega &= b(\dot{\mathbf{R}}\ddot{\mathbf{C}}\dot{\mathbf{R}}^{\dagger}) \\ &= b(\dot{\mathbf{X}}^{\dagger}\ddot{\mathbf{C}}\dot{\mathbf{X}}) \\ &= b(\dot{\mathbf{H}}_0^{\dagger}\ddot{\mathbf{C}}^{-1}\dot{\mathbf{H}}_0)^{-1} \\ &= b/\rho_{max}. \end{aligned}$$

The factor $\dot{\mathbf{H}}_0^{\dagger}\ddot{\mathbf{C}}^{-1}\dot{\mathbf{H}}_0$ is unimportant because it is considered to be perfectly known *in the directivity concept*. Since the signal and noise are uncorrelated, the total spectral density on the output is $d + b/\rho_{max}$. The error in the spectral analysis of the signal is b/ρ_{max} and may be negligible for a large value of ρ_{max}, i.e., a large array gain. Doing better requires some knowledge of b in order to "identify" d, the only parameter which remains unidentified in the directivity concept.

2.8

The real problem is to get an estimate for b. Herein arises the notion of "signal-free reference." The possibility of building a combination of sensors from which only the noise would appear is, at first sight, relatively easy. If we choose a set of filters represented by a row vector $\dot{\mathbf{V}}$ such that $\dot{\mathbf{V}}\dot{\mathbf{H}}_0 = 0$, then the signal with a wavefront $\dot{\mathbf{H}}_0$ will not appear at the output. The omnidirectional noise will appear at the beamformer output with spectral density $b(\dot{\mathbf{V}}\ddot{\mathbf{C}}\dot{\mathbf{V}}^{\dagger})$ which can be measured and from which b can be determined. In principle the choice of $\dot{\mathbf{V}}$ is arbitrary within the $(N-1)$ dimensional subspace orthogonal to $\dot{\mathbf{H}}_0$.

In the current classical preformed-beam techniques, the role of the "signal-free" combination is approximately played by the mean value of the outputs on beams adjacent to or close by the one being tested. The fact that only neighboring beams are considered clearly reveals our uneasiness about the omnidirectionality assumption.

The drawback with the previous procedure is that there may be something other than the signal ω_0 and the omnidirectional noise present. In particular, there may be other sources which are not well described in the directivity concept as unwanted omnidirectional noise. These other sources can add unknown contributions to both combinations $\dot{\mathbf{R}}_0$ and $\dot{\mathbf{V}}$, and can be labeled as "unwanted signals."

Concerning the presence of unwanted signals arriving at the array as plane waves which are uncorrelated with each other, with the ω_0 source, and with the omnidirectional noise, two situations can arise:

1. In the first situation, the direction of each unwanted signal is known, so that we have knowledge of

$$\begin{array}{lll} \omega_1 & \text{and therefore} & \dot{\mathbf{H}}_1 \\ \omega_2 & \text{and therefore} & \dot{\mathbf{H}}_2 \\ \vdots & & \vdots \\ \omega_P & \text{and therefore} & \dot{\mathbf{H}}_P. \end{array}$$

One must then seek a "source-free" reference for the omnidirectional noise which has a vector $\dot{\mathbf{V}}$ orthogonal to all the vectors

$$\dot{\mathbf{H}}_0, \dot{\mathbf{H}}_1, \ldots, \dot{\mathbf{H}}_p, \ldots, \dot{\mathbf{H}}_P.$$

For this to be possible, there must exist a subspace "free of sources." Thus, N, the number of sensors in the array, must be larger than $P + 1$, the number of sources present. This is but the first illustration of how the array has to be large enough to surmount the complexity of the source system. This topic will be discussed further in later sections.

If the condition $N > P + 1$ is fulfilled, it is possible to remove the omnidirectional noise spectral density from the outputs of the combinations derived from the $\dot{\mathbf{H}}_p$ according to (2.11). At the $\dot{\mathbf{H}}_0$ output

we will still have spectral contributions due to the unwanted signals. Spectral analysis will yield

$$\begin{aligned} A_0 &= d + d_1 |\bar{\mathbf{R}}_0 \vec{\mathbf{H}}_1|^2 + \cdots + d_P |\bar{\mathbf{R}}_0 \vec{\mathbf{H}}_P|^2 \quad \text{at the } \vec{\mathbf{H}}_0 \text{ output,} \\ A_1 &= d |\bar{\mathbf{R}}_1 \vec{\mathbf{H}}_0|^2 + d_1 + \cdots + d_P |\bar{\mathbf{R}}_1 \vec{\mathbf{H}}_P|^2 \quad \text{at the } \vec{\mathbf{H}}_1 \text{ output,} \\ &\vdots \\ A_P &= d |\bar{\mathbf{R}}_P \vec{\mathbf{H}}_0|^2 + \cdots + d_P \quad \text{at the } \vec{\mathbf{H}}_P \text{ output.} \end{aligned}$$

From this system of $(P+1)$ equations for the $(P+1)$ unknown spectral densities one can identify all the sources in terms of their individual spectral densities.

2. In the second situation, the directions of the unwanted sources are unknown, but the number of sources is expected to be much less than N. It is then still possible, in principle, to remove the omnidirectional noise from the $\vec{\mathbf{H}}_0$ output. By sweeping with a variable combination $\vec{\mathbf{V}}$ all the subspace orthogonal to $\vec{\mathbf{H}}_0$, we can determine a cone in which the noise is both minimum and constant at the output of $\vec{\mathbf{V}}$. This value is qualified to represent the omnidirectional noise. Such a process is certainly painful and will be used reluctantly, because the strict omnidirectional noise assumption can not be trusted to such an extent. Anyway there is now no way to eliminate the contributions from a random distribution of other sources. Identification of the spectral density d is hazardous if not impossible.

In the practical implementation of classical beamforming, where every wavefront is taken to be a plane wave, one seeks to build the array large enough so that with its many narrow beams it can minimize the scalars $(\bar{\mathbf{R}}_p \vec{\mathbf{H}}_q)$ and limit the interference between sources. Within the limited framework provided by the directivity concept there is little more that can be done.

2.9

The discussion of the directivity concept can be summarized as follows:

1. Only its expression as a receiving array gain need be considered here.

2. The description it employs of signal and noise structure is over-simplified compared to that in the real world. Thus, this description relies on extremely stringent assumptions.

3. The status of waves differing from the one carrying the signal is somewhat vague and the possible presence of unwanted signals considerably complicates the identification of the wanted signal.

4. The role played by the correlation matrix is modest. Actually, because the function $D(x, f)$ is usually directly accessible, the appearance of the cross-spectral density matrix in the presentation in section 2.4 was unnecessary and somewhat forced. The only reason for introducing it was to present the directivity concept in a manner which will be essential in later sections.

As a formal and final remark we note that once $\mathbf{H}$ had been chosen to be a column matrix, it was quite natural to choose $\mathbf{R}$ to be a row matrix. Because it is desirable to handle only vectors of the same species, however, we can choose the filter vector also to be a column vector. With the convention that the frequency responses appearing in the filter column vector are taken to be the complex conjugates of the actual frequency responses, the transpose conjugate of the filter column vector will be identical to the row vector with which we have thus far dealt. Henceforth we shall deal only with column vectors.

Chapter 3

ADAPTIVE SIGNAL PROCESSING WITH A KNOWN WAVEFRONT

3.1

We first have to comment about stationarity and adaptivity. Consider a *remote point like source transmitting a random signal $S(t)$. This signal* reaches a set of N sensors after passage through a propagation medium. The geometry of the array is known and can be referred to in terms of coordinates. Usually the spacing of the sensors is about half a wavelength of the highest expected frequency. The outputs of these sensors are N voltages exclusively caused by $S(t)$ as modified by the medium. So far no other source or noise interference is presumed to be present.

The relation between the N sensor outputs can then be described in terms of a correlation matrix in the spatio-temporal domain, or of a cross-spectral density matrix (CSDM) in the spatio-spectral domain. Both descriptions are equivalent; the second will be preferred here for its simpler mathematical symbolism. Although the CSDM is used, we are primarily interested in large bandwidths and not pure sine waves. From a strict point of view, correlation implies stationarity and neither the signal $S(t)$ nor the medium usually meet this criterion. Nevertheless the *limited time quasi-stationarity assumption* is the cornerstone of adaptive processing in both spatio temporal and spatio-spectral domains.

More precisely we shall assume that the fluctuations of the medium are slow enough to be considered as quasi-stationary over a time T, which for the ocean is estimated to be of the order of one second. Concerning the signal $S(t)$, we also assume the same time T of quasi-stationarity. Furthermore we shall be concerned with signals of more than a thousand Hertz bandwidth W. With such a large time-bandwidth product ($WT > 1000$), the problem of estimating the second order statistics should *not be too critical*. Anyway such estimation methods have been the focus of much scientific literature [1] [13] and improvements can still be reasonably expected.

The present paper assumes the validity of these estimation methods. Under these conditions the elements of a CSDM can be estimated with little significant error. It is then pointless, and frequently unwise, to impose an artificial relationship, such as the Toeplitz assumption, upon the experimental values. This position being adopted, two remarks must be made.

1. All the estimation methods commonly in use today implicitly rely on the *existence of a stationary process which is estimated* during a limited time interval. Actually the real problem is how best to fit this stationary process to a non-stationary data sequence. Although this distinction may seem slight, it could be the basis for refining the estimation methods.

2. Every estimation method relies on assumptions about the process influencing the data outside the available sequence. It seems reasonable that a "good" method should be one with few assumptions. This concept will be used as a guideline in this paper.

In short, we will only consider spectral and cross-spectral densities which are valid for a limited time and experimentally renewed every T seconds or faster, such as the DELTIC (Delay Line Time Compression) procedure[1] [11].

[1] This almost forgotten procedure, which was originally implemented using delay lines, can be transposed to computer algorithms. It provides the maximum amount of overlapping data (maximum adaptivity) compatible with the sampling rate, at the cost of some computer speed multiplied by a factor obtained from the time-bandwidth product.

3.2

We must now be quite clear about the source vector and wavefront. Within the framework of the previous quasi-stationarity concept, let us recall the single source transmitting $S(t)$. The transfer function between this source and the *output* of the n^{th} sensor can be represented by a linear filter with a complex frequency response $\phi_n(f)$ which is a component of the N-dimensional complex vector $\vec{\Phi}(f)$.

If $d_t(f)$ is the *true* spectral density at the source, the cross-spectral density between the output of sensors n and m is given by a well known expression already presented in

$$q_{nm} = d_t(f)\phi_n(f)\phi_m^*(f).$$

To disencumber the notation, the frequency will henceforth be treated as an implicit variable except when the contrary is specifically stated. With this change in notation,

$$q_{nm} = d_t\phi_n\phi_m^* \tag{3.1}$$

is the n^{th} row, m^{th} column element of the CSDM linked to the source. By convention Φ is given the status of a *column* matrix, then the CSDM is

$$\vec{\mathbf{Q}} = d_t\vec{\Phi}\vec{\Phi}^\dagger. \tag{3.2}$$

We then define the "source vector" as

$$\vec{\mathbf{S}} = d_t^{\frac{1}{2}}\vec{\Phi}, \tag{3.3}$$

which allows the CSDM to be written as a dyadic product of the vector $\vec{\mathbf{S}}$,

$$\vec{\mathbf{Q}} = \vec{\mathbf{S}}\vec{\mathbf{S}}^\dagger. \tag{3.4}$$

We also find it convenient to introduce the concept of "wavefront" as the *normalized* vector (and also column matrix)

$$\vec{\mathbf{F}} = \frac{\vec{\Phi}}{|\vec{\Phi}|}$$

where the scalar

$$|\Phi| = \left(\Phi^\dagger\Phi\right)^{\frac{1}{2}} \tag{3.5}$$

and obviously $\vec{\mathbf{F}}^{\dagger}\vec{\mathbf{F}} = 1$. Vector $\vec{\mathbf{F}}$ is *not the physical wavefront* but only its *transfer* through the *medium and array sensors*. For the sake of lexical simplicity let us accept this slight discrepancy with orthodox physics. The wavefront is then associated with some particular source position and associated with it is a fixed set of filters.

The spectral density d_t of the source is unknown since it is the object of identification. The vector $\vec{\Phi}$ (of filters ϕ_n) is also unknown since we dismiss any knowledge of the medium in this paper. What can be measured at the output of the sensors are the components of the source vector

$$\begin{aligned}\vec{\mathbf{S}} &= d_t^{\frac{1}{2}}\vec{\Phi} \\ &= d_t^{\frac{1}{2}}|\vec{\Phi}|\,\vec{\mathbf{F}}\end{aligned}$$

with their relative amplitude and phases. But with the absence of any knowledge about the source, it will never be possible to isolate $d_t^{\frac{1}{2}}$ from the unknown $|\vec{\Phi}|$. So introduction of the wavefront vector is justified, and the array is only able to yield an *apparent* spectral density, d, given by

$$d^{\frac{1}{2}} = d_t^{\frac{1}{2}}|\vec{\Phi}|. \tag{3.6}$$

This relation reflects the ability of interference, for example between coherent multipaths, to alter the spectral density of the source as it is seen through the array. Finally, the expression for the source vector is

$$\vec{\mathbf{S}} = d^{\frac{1}{2}}\vec{\mathbf{F}} \tag{3.7}$$

with $\vec{\mathbf{F}}^{\dagger}\vec{\mathbf{F}} = 1$ and $\vec{\mathbf{S}}^{\dagger}\vec{\mathbf{S}} = d$.

In this section, the wavefront vector is supposed to be *known a priori*, even if it does not correspond to a plane or spherical wave. The goal of identification is the analysis of the only accessible spectral density, the apparent one, d.

3.3

At this point we combine the sensor outputs to obtain an estimate of the apparent spectral density using concepts mentioned previously in chapter 2.

This is relatively easy since the wavefront vector is known by its N complex components f_n. We then build a set of filters with the complex conjugate responses f_n^*. Adding the output of these N parallel filters produces the global filter

$$\sum_n f_n f_n^* = \vec{\mathbf{F}}^\dagger \vec{\mathbf{F}} = 1 \tag{3.8}$$

The frequency response of this all-pass filter yields the apparent spectral density of the signal for the purposes of analysis and identification. This will be referred to as the "main output."

Incidentally the filters f_n may not be causal. When this occurs they must be built with the same additional time delay, in order to translate all the impulse responses equally within the domain $t \geq 0$. Also, when a set of filters is claimed to be represented by a column matrix $\vec{\mathbf{F}}$, it is, in fact, the row matrix $\vec{\mathbf{F}}^\dagger$ with complex conjugate frequency components of $\vec{\mathbf{F}}$. With this convention, it is not necessary (as in chapter 2) to use different symbols for a wavefront and a set of filters matched to that wavefront (see section2.9).

Dropping the fiction of a single source, the presence of some unwanted noise in the main output will now be considered. Contrary to chapter 2, the spatial structure of this noise is unknown. Indeed, *it does not need to be known*. The only requirement a source must meet to be considered noise is a lack of correlation with $S(t)$. In general, the source may be either spatially concentrated, such as a jammer with an allowable wavefront vector which is different from $\vec{\mathbf{F}}$, or diffuse such as background shipping without a distinct wavefront. The structure of this global noise will be adaptively learned by measures on the sensor outputs and estimation of the CSDM. While the set of filters $\vec{\mathbf{F}}$ is not adaptive, everything else from now on will be, with the time constant T.

3.4

Unfortunately, as in chapter 2, we can never know whether the estimated CSDM corresponds to signal plus noise or only to noise. Nevertheless it is possible to get an indication of the noise by using knowledge of $\vec{\mathbf{F}}$. Since the set of filters with components l_n represented by a vector $\vec{\mathbf{L}}$ orthogonal

to $\vec{\mathbf{F}}$ eliminates the signal,

$$\vec{\mathbf{L}}^{\dagger}\vec{\mathbf{F}} = 0 \tag{3.9}$$

or equivalently

$$\sum_{n} l_n^{*} f_n = 0, \tag{3.10}$$

the output of the combination of sensors represented by $\vec{\mathbf{L}}$ will be a "signal-free reference" (SFR).

Within the N-dimensional space of the complex column vectors which span the CSDM, the entire $(N-1)$-dimensional subspace orthogonal to $\vec{\mathbf{F}}$ is available to vectors of the $\vec{\mathbf{L}}$ type. The full exploitation of signal-free information[2] is then dependent on this subspace which may be spanned by a set of linearly independent vectors $\vec{\mathbf{L}}_j$, $(1 \leq j \leq N-1)$ orthogonal to $\vec{\mathbf{F}}$. Since we may arbitrarily choose the filters subject to the constraint

$$\vec{\mathbf{L}}_j^{\dagger}\vec{\mathbf{F}} = 0 \tag{3.11}$$

we select a set which satisfies the condition of inter-orthogonality

$$\vec{\mathbf{L}}_j^{\dagger}\vec{\mathbf{L}}_k = 0, \quad j \neq k. \tag{3.12}$$

Additionally we can further restrict the filter vectors by requiring *noncorrelation* between the output of combination $\vec{\mathbf{L}}_j$ and that of combination $\vec{\mathbf{L}}_k$,

$$\vec{\mathbf{L}}_j^{\dagger}\vec{\mathbf{C}}\vec{\mathbf{L}}_k = 0, \quad j \neq k. \tag{3.13}$$

This point is discussed in more detail in appendix D.

The vectors $\vec{\mathbf{L}}$ must be derived from (3.11) (3.12) and (3.13). One might argue that $\vec{\mathbf{C}}$, the CSDM of the noise alone, is not experimentally estimable. However because of (3.11), (3.13) can be rewritten as

$$\vec{\mathbf{L}}_j^{\dagger}\left[\vec{\mathbf{C}} + d\vec{\mathbf{F}}\vec{\mathbf{F}}^{\dagger}\right]\vec{\mathbf{L}}_k = 0, \quad j \neq k, \tag{3.14}$$

where the matrix in brackets is a signal plus noise CSDM. As a result we are allowed to introduce the CSDM estimated on the sensors into (3.13), *regardless of signal presence or absence*.

[2] Contrary to current assumptions, there certainly is never a noise alone situation at the output of the $\vec{\mathbf{F}}$ combination. This is the usual predicament in long range passive surveillance.

Let us now see under what conditions (3.11), (3.12) and (3.13) can yield a unique solution for the filter coefficients $\vec{\mathbf{L}}_j$. None of these equations can determine the arbitrary complex scale factor of $\vec{\mathbf{L}}_j$. So the actual number of unknown scalars to derive is only $N-1$. Since there are $N-1$ vectors $\vec{\mathbf{L}}_j$ subject to this system of equations, there are $(N-1)^2$ complex scalars to be derived.

Now (3.11) corresponds to $N-1$ scalar equations while (3.12) and (3.13) each correspond to $(N-1)(N-2)/2$ scalar equations obtained as j steps from 1 to $N-1$ and k steps from 1 to $j-1$. Totaling the number of scalar equations we get $(N-1)+2(N-1)(N-2)/2=(N-1)^2$ which is the number of equations required to determine all the complex scalars uniquely.

We then have a saturated system of equations from which the vectors $\vec{\mathbf{L}}_j$ can be derived. Although these filter coefficients are completely independent of signal presence or absence, they are adaptive to the noise. Thus it is possible to transform the original set of N sensor outputs to another system for which the "main output" lets signal through as an all-pass filter (unaltered signal with the spectral density d), and the other $(N-1)$ outputs are uncorrelated with the signal and each other but possibly correlated with the noise in the main output. From now on $\vec{\mathbf{C}}$ will denote this transformed CSDM.

3.5

The above possibility is so important that it is worth further illustration by a different approach. Let us first require the vectors $\vec{\mathbf{L}}_j$ to be orthogonal only to $\vec{\mathbf{F}}$ but not to each other, and to be distinct so a full description of the noise subspace is possible. In particular we can easily build the non-adaptive vectors:

$$\vec{\mathbf{L}}_1 \begin{pmatrix} -f_2^* \\ f_1^* \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{\mathbf{L}}_j = \begin{pmatrix} -f_j^* \\ 0 \\ 0 \\ \vdots \\ f_1^* \\ \vdots \\ 0 \end{pmatrix} \leftarrow j^{\text{th}}\text{ row} \quad \cdots \quad \vec{\mathbf{L}}_{N-1} = \begin{pmatrix} -f_{N-1}^* \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ f_1^* \end{pmatrix}; \qquad (3.15)$$

which are distinct since they use different pairs of sensors and which satisfy $\dot{\mathbf{L}}_j^{\dagger}\dot{\mathbf{F}} = 0$.

These $\dot{\mathbf{L}}_j$ combinations yield correlated SFRs, however we wish to retain the signal free-information in a set of uncorrelated SFRs. To do this we linearly transform a set of correlated SFRs to a set of uncorrelated SFRs which are their "orthogonal images." The method adaptively estimates the eigenvectors of the noise CSDM subspace orthogonal to $\dot{\mathbf{F}}$. Then, every eigenvector is used to form a linear combination of the original SFRs. Because of the particular properties of eigenvectors, the new SFRs are no longer correlated. This process is the physical expression of the diagonalization of a matrix which is discussed in appendix D.

With respect to this new set of basis vectors, the CSDM has a specific aspect involving diagonal elements which represent spectral densities, and off-diagonal elements in the first row and column which represent cross-spectral densities with the main output; all other off-diagonal elements are zero because the SFRs have been mutually decorrelated,

$$\ddot{\mathbf{C}} = \begin{pmatrix} c_{00} & c_{01} & c_{02} & \cdots & c_{0N'} \\ c_{10} & c_{11} & 0 & \cdots & 0 \\ c_{20} & 0 & c_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{N'0} & 0 & 0 & \cdots & c_{N'N'} \end{pmatrix}, \quad \text{where } N' = N - 1. \tag{3.16}$$

All estimates of the elements are signal-free except c_{00}. Indeed, any cross-spectral density between the main output and a SFR is signal-free, even when signal is present, since signal is not correlated with any noise, nor with any linear combination of noises such as a SFR. Any cross-spectral density between two SFRs is zero since the noise has been decorrelated.

We reach the situation illustrated in figure 3.1 in which a main output with a possible signal of spectral density d and a noise B_0 with a spectral density c_{00} is combined to produce the final output with the uncorrelated SFR's, B_1 to $B_{N'}$ having spectral densities c_{11} to $c_{N'N'}$. The spectral densities c_{nn} can be adaptively measured on the corresponding outputs except c_{00} which may be "polluted" by the signal presence.

We now build a vector $\dot{\mathbf{V}}$ representing a set of filters which is an all pass filter for the signal and which yields a minimized noise spectral density. This vector has $v_0 = 1$ for the first component and combinations of the main

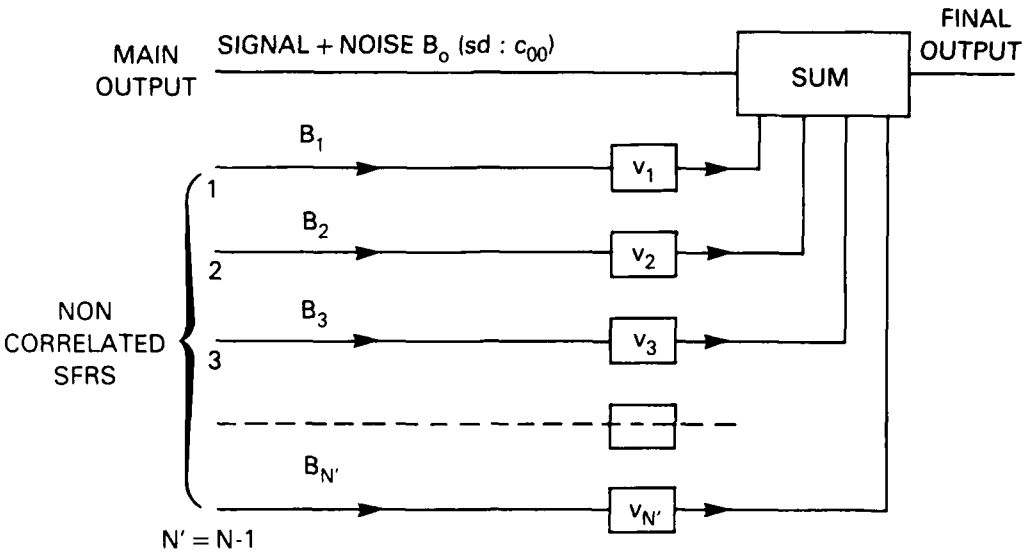


Figure 3.1: Simplified Processor with $(N-1)$ Decorrelated Signal-Free References.

output and SFRs for the remaining components,

$$\vec{\mathbf{V}}^{\mathrm{T}} = [1, v_1, v_2, \ldots v_{N'}] \tag{3.17}$$

The resulting spectral density projected on $\vec{\mathbf{V}}$ is given by the quadratic form $\vec{\mathbf{V}}^{\dagger}\vec{\mathbf{C}}\vec{\mathbf{V}}$. The constraint $v_0 = 1$ on vector $\vec{\mathbf{V}}$ can be expressed as follows. Let us consider an N-dimensional vector $\vec{\Pi}$ with all its components null except the first one, $\vec{\Pi}^{\mathrm{T}} = [1, 0, 0, \ldots 0]$. The constraint can then be written as

$$\vec{\Pi}^{\dagger}\vec{\mathbf{V}} = \vec{\mathbf{V}}^{\dagger}\vec{\Pi} = 1,$$

which are particular cases of

$$\left|\vec{\mathbf{V}}^{\dagger}\vec{\Pi}\right| = 1. \tag{3.18}$$

We can then vary vector $\vec{\mathbf{V}}$ so as to minimize $\vec{\mathbf{V}}^{\dagger}\vec{\mathbf{C}}\vec{\mathbf{V}}$ which is equivalent to maximizing

$$\rho \qquad \frac{\left|\vec{\mathbf{V}}^{\dagger}\vec{\Pi}\right|^2}{\vec{\mathbf{V}}^{\dagger}\vec{\mathbf{C}}\vec{\mathbf{V}}} \tag{3.19}$$

The solution is derived as in section 2.6, with:

$$\begin{aligned} \vec{\mathbf{X}} &\quad \text{replaced by} \quad \vec{\mathbf{V}}, \\ \vec{\mathbf{H}}_0 &\quad \text{replaced by} \quad \vec{\Pi}, \\ \vec{\mathbf{Y}} &\quad \text{replaced by} \quad \vec{\mathbf{M}}\vec{\mathbf{V}}, \\ \vec{\mathbf{P}}_0 &\quad \text{replaced by} \quad \left(\vec{\mathbf{M}}^{\dagger}\right)^{-1}\vec{\Pi}. \end{aligned}$$

Equation 1.14 can be written as

$$\vec{\mathbf{C}}\vec{\mathbf{V}} = \alpha\vec{\Pi}, \tag{3.20}$$

where α is an arbitrary scalar. What matters is that (3.20) expresses the fact that the components of $\vec{\mathbf{C}}\vec{\mathbf{V}}$ are nullified except for the first one. Using (3.16) and (3.17) we now write,

$$c_{j0} + v_j c_{jj} = 0, \quad j = 1, N'. \tag{3.21}$$

Two remarks must be made: first, the spectral density c_{00} does not appear in (3.21), hence it is also absent from the components v_j. Thus vector $\vec{\mathbf{V}}$ represents signal-free information. Second, every component of $\vec{\mathbf{V}}$ (except $v_0 = 1$) is dependent *only* on the spectral density of the corresponding SFR and the cross-spectral density of this SFR with the noise on the main output. This last remark is intuitively predictable from the non-correlation of the SFRs. The entire process is *a repetition of independent processes* for each SFR. The filter to be put on the j^{th} SFR is the complex conjugate of the component v_j,

$$v_j^* = -\left(\frac{c_{j0}}{c_{jj}}\right)^* = \frac{c_{0j}}{c_{jj}}. \tag{3.22}$$

Although we might now give a formal expression of the minimized noise $\vec{\mathbf{V}}^{\dagger}\vec{\mathbf{C}}\vec{\mathbf{V}}$ with the vector $\vec{\mathbf{V}}$ we have just determined, it seems more fruitful to look into the physical significance of such a result.

3.6

Let us consider separately the main output and the first SFR. Filtering the latter with $-c_{01}/c_{11}$ and adding the result to the main output is equivalent

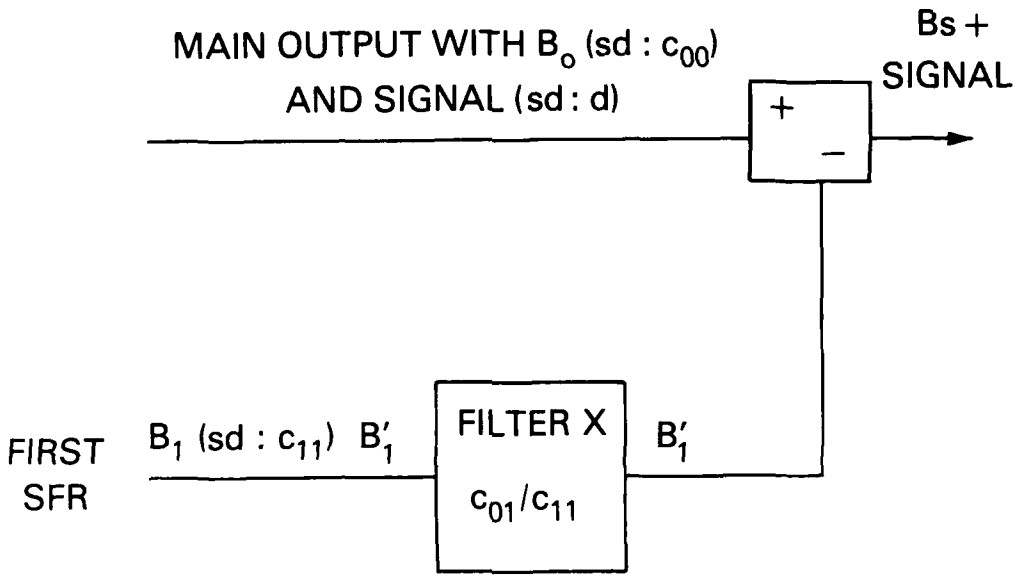


Figure 3.2: Processor with a Single SFR Noise Subtracter.

to filtering the SFR with c_{01}/c_{11} and subtracting the result (see figure 3.2). Let us now proceed to design a filter, x (also shown in figure 3.2), which minimizes the spectral density of the noise in the output, $B_s = B_0 - B_1'$. Hence using the rule given in appendix A with the notation sd $\equiv$ spectral density and csd $\equiv$ cross-spectral density:

$$\begin{aligned} \mathrm{sd}[B_s] &= \mathrm{csd}[B_s B_s] \\ &= \mathrm{csd}[(B_0 - B_1')(B_0 - B_1')] \\ &= \mathrm{sd}[B_0] - \mathrm{csd}[B_1' B_0] - \mathrm{csd}[B_0 B_1'] + \mathrm{sd}[B_1'] \end{aligned}$$

But:

$$\begin{aligned} \mathrm{csd}[B_1' B_0] &= c_{10} x \\ \mathrm{csd}[B_0 B_1'] &= c_{01} x^* \\ \mathrm{sd}[B_1'] &= c_{11} |x|^2 = c_{11} x x^* \end{aligned}$$

So that

$$\mathrm{sd}[B_s] = c_{00} - c_{10} x - c_{01} x^* + c_{11} x x^* \tag{3.23}$$

Varying x to get a minimum, we must have for any value of the increment δx

$$\delta x^* [c_{11} x - c_{01}] + \delta x [c_{11} x^* - c_{10}] = 0$$

hence

$$c_{11}x - c_{01} = 0 \tag{3.24}$$

which again justifies

$$x = \frac{c_{01}}{c_{11}}.$$

It is remarkable that the left side of (3.24) is precisely the cross-spectral density of B_s with B_1. Indeed looking at figure 3.2:

$$\begin{aligned} B_s &= B_0 - B_1' \\ \mathrm{csd}[B_s B_1] &= \mathrm{csd}[B_0 B_1] - \mathrm{csd}[B_1' B_1] \\ \mathrm{csd}[B_0 B_1] &= c_{01} \\ \mathrm{csd}[B_1' B_1] &= (\mathrm{csd}[B_1 B_1'])^* \\ &\quad (c_{11}x^*)^* \\ &= c_{11}x \end{aligned}$$

hence

$$\mathrm{csd}[B_s B_1] = c_{01} - c_{11}x$$

which is precisely the quantity nullified in (3.24) to minimize the output noise. It so happens that *the same filter which provides the minimum of output noise, also nullifies the correlation between the output and the SFR involved.*

In very qualitative (even hardly scientific) terms a physical interpretation could be presented as follows. The filter c_{01}/c_{11} draws from the SFR what could be considered as the best "likeness" of B_0 so that the subtraction of this likeness yields a minimum of residual noise. As a result, the "likeness" of the SFR has been removed from the residual noise remaining in B_s, in other words B_s finds itself decorrelated from the SFR.

The spectral density of B_s as derived from (3.23) and (3.24) is

$$\mathrm{sd}[B_s] = c_{00} - \frac{|c_{01}|^2}{c_{11}} \tag{3.25}$$

Since the spectral density of B_1' is $|c_{01}|^2 / c_{11}$, we can see that for this particular operation it happens that *subtracting two noises is equivalent to subtracting their spectral densities;* this is far from a general rule and may, at first glance, even be looked upon as paradoxical.

3.7

The repetition of the same process with the other SFRs is obvious. With j stepping from 1 to $(N-1)$, we put a filter, c_{0j}/c_{jj}, on the j^{th} SFR, B_j, and subtract the result B'_j from the main output to obtain the set of spectral density reductions, $|c_{0j}|^2/c_{jj}$. When all the SFRs are exhausted, the output of the whole process is uncorrelated with each SFR and possesses the minimized spectral density

$$c_{\min} = c_{00} - \sum_{j=1}^{N-1} \frac{|c_{0j}|^2}{c_{jj}}. \tag{3.26}$$

It is easy to check that $c_{\min}$ is positive and precisely the value of $\vec{\mathbf{V}}^\dagger \vec{\mathbf{C}} \vec{\mathbf{V}}$ (from section 3.5) when $\vec{\mathbf{V}}$ is the solution of (3.22). Since the signal is not correlated with the SFRs, the unaltered signal spectral density d may be additively superposed on $c_{\min}$. While the sum on the right side of (3.26) only involves signal free estimates, c_{00} may or may not contain signal. As a result we still can not say for sure whether the spectral analysis of the final output yields $c_{min} + d$ or c_{min} alone.

Although we are in a common situation with a signal and a minimized noise under the constraint of unaltered signal (Wiener), we are close to achieving something better. Indeed *only one scalar relation* between the elements of the noise alone matrix $\vec{\mathbf{C}}$ would provide a way of computing noise alone, c_{00}, as a function of the other elements. We have exhausted the full array power and now need some *a priori* information. Incidentally, it is easy to check that the wanted relation *is not* provided by the computation of c and of the first component of vector $\vec{\mathbf{C}}\vec{\mathbf{V}}$ in (3.20). So if any bit of *reliable* information is available about the original noise, it can be transposed to the SFRs and provide a relation. That would give us a "super-resolution" system where the second order statistics of the unwanted noise could be eliminated, and the spectral density of the signal could be derived from the spectral analysis of the final output with, in principle, no error.

3.8

If no assumption is suggested by the physics of the situation, we can still try to illustrate how the "super-resolution system," yielding *no error* in

signal identification, could be considered an "asymptotically large array assumption." The original CSDM, $\ddot{\Gamma}$, taken directly behind the sensors in a truly operational situation cannot show an *exactly null* eigenvalue. But the smallest may be very small as compared to the largest one. In fact a test that the array has enough sensors to "surmount the complexity" of the acoustic field should appear when looking at the set of eigenvalues. This set should exhibit a large span of values. Intuitively the smallest should decrease when the number of sensors, N, increases. The validity of this "asymptotically large array assumption" is dependent *not only on* N being much larger than the number of sources, but also on the medium complexity and the structure of the diffuse noise sources (see chapter 11).

Now for any CSDM with a non-zero smallest eigenvalue ϵ_{min}, we can write

$$\ddot{\Gamma}' = \ddot{\Gamma} - \epsilon_{min}\ddot{\mathbf{I}}_N,$$

where $\ddot{\mathbf{I}}_N$ stands for the unit matrix of order N. Then $\ddot{\Gamma}'$ is a "singular" matrix with one zero eigenvalue or more. A well known physical interpretation can be proposed in which ϵ_{min} is the spectral density of a fraction of noise which is uncorrelated between sensors and has the same spectral density on each sensor. This interpretation of ϵ_{min} is typical of a sort of background noise of separate but identical preampliliers. But such an electric noise is known to lie far below the minimum acoustic level in most arrays. Such an interpretation of ϵ_{min} does not mean that it is the *physical reality* for any CSDM, $\ddot{\Gamma}$, but as the array grows in size and the span of the eigenvalues spreads, it becomes more and more likely that the decreasing ϵ_{min} could really represent this kind of noise. It would be still more credible if two or three eigenvalues were found with almost the same minimum value; and of course, a minimum practically negligible as compared to the largest eigenvalues. The "asymptotically large array assumption" consists in stating that *this sort of noise cannot exist acoustically* and therefore the ϵ_{min} yielded by the original array can be simply neglected.

In the frame of such an assumption $\ddot{\Gamma}$ becomes, like $\ddot{\Gamma}'$, a "singular" CSDM. Since the main output and the SFRs are derived linearly from the sensor outputs, the CSDM $\ddot{\mathbf{C}}$ is also singular, which means

$$\det(\ddot{\mathbf{C}}) = 0, \tag{3.2c}$$

where "det" denotes the determinant of the matrix. To accept (3.2c) is to deny the existence of this very stringently determined "background"

acoustic-noise on the main output and the SFRs. It is easy to check that (3.26) also reads

$$c_{\min} = \frac{\det(\ddot{\mathbf{C}})}{c_{11}c_{22}\cdots c_{N'N'}}, \tag{3.28}$$

so that accepting (3.27) is simply accepting that there is no noise on the final output and we have got a "super-resolution" system. This can only be credible as a limiting representation of infinitesimal "background" acoustic noise if $\ddot{\mathbf{C}}$ possesses a large span of eigenvalues and if the minimum value is reasonably negligible.

3.9

Let us now reconsider figure 3.2 which shows a "noise subtractor." This terminology is generally associated with "noise canceling" techniques, where the unwanted noise on the final output is really nullified. In fact noise canceling is the same device as in figure 3.2 but used in a particularly favorable situation where B_1 and B_0 are "fully correlated." It means that both noises coming from one source through linear filtering, B_0 can be seen as a noise produced from B_1 through a linear filter ϕ. We then have

$$\begin{aligned} \mathrm{csd}[B_1 B_0] &= c_{11}\phi^* = c_{10} \\ \mathrm{sd}[B_0] &= c_{11}|\phi|^2 = c_{00} \end{aligned}$$

and therefore

$$c_{00}c_{11} = |c_{10}|^2 \tag{3.29}$$

for any possible ϕ. This is the usual test of "full correlation." Equation (3.25) then yields $\mathrm{sd}[B_s] = 0$ and the noise subtractor is as nearly perfect as full correlation is achieved.

It is time now to complete figure 3.2 with the processors it implies. This is done in figure 3.3. The two time delay devices (or algorithms) are there to ensure that the subtraction is really achieved with the very same sequences of B_0 and B_1 which participated in the building of the filter with a time constant T. This device, or the algorithmic equivalent of figure 3.3, is known in the French scientific literature as a "correlo-filtre" and has been extensively studied and improved in the past fifteen years [15–17].

A particular feature of the correlo-filter, in fact a degree of freedom for the processor, is that any preliminary linear filtering of the SFR does not

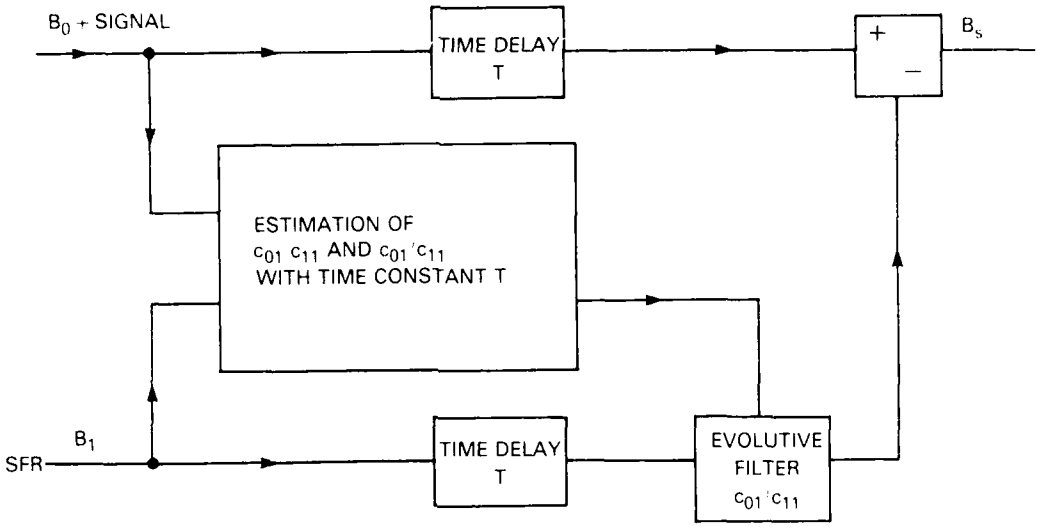


Figure 3.3: Adaptive Processor with a Single SFR Correlo-Filter.

change the spectral density of B_s. This is why, in its present form, we have *preferred* to first "self-whiten" the SFR, B_1, with an open or closed loop algorithm and then proceed with a normalized value of the spectral density, $c_{11} = c$, which is *not frequency dependent* in the useful bandwidth.

A self-whitening filter must have a frequency response W_1 such that $|W_1|^2 = c \cdot c_{11}$. The phase law of this filter is free which makes it easier to build, and, as mentioned above, does not affect the result. Figure 3.3 is then modified as shown in figure 3.4, where the filtering of the whitened SFR B_1 can be conceived in the time domain as its convolution with the cross-correlation function of B_0 with B_1 and the essential scale factor $1 \cdot c$.

If all the other SFRs are whitened at the same spectral density c and if we name correlo-filter (CF) that which is framed by the dotted line in figure 3.4, the processing from the sensor outputs can be sketched as in figure 3.5. This figure could be presented as an open loop solution[3]to the original problem of getting the signal spectral density with a minimum of unwanted noise and, for a large array, with practically no noise at all. Nevertheless it is not intended here to propose such an open loop algorithm for operational

[3]The process of the SFR diagonalization, the building of orthogonal images, can be achieved with a "cascading" sequence of CFs. So that an open loop solution is conceptually built with the CF as a standard module, the universal brick of spatial processing [15].

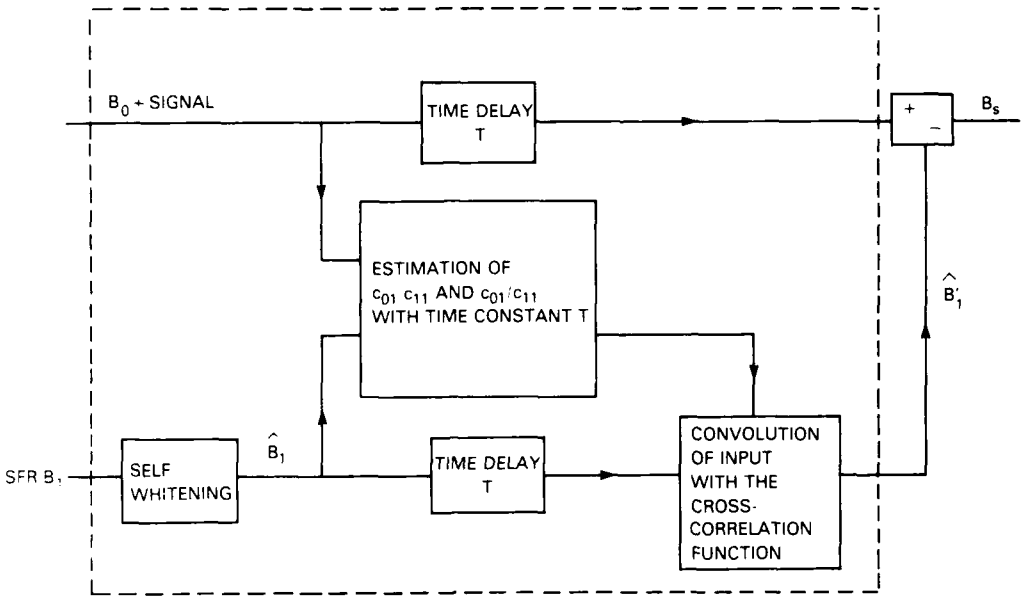


Figure 3.4: Correlo-Filter in the Time Domain (with Self Whitening of the SFR).

practice, but rather as a tool to illustrate the physical meaning of the successive steps and to allow experimental checking of the specific properties described at each step. Still, any practical solution has to preserve the fundamental property of a processor which is independent of the signal presence or absence.

3.10

A serious weakness in adaptive processing of known signal wavefronts lies precisely in the reliability of the a priori information. When the signal wavefront happens to slightly differ from the vector $\vec{\mathbf{F}}$, which it is supposed to equal, signal components will appear in the SFRs whose vectors $\vec{\mathbf{L}}$ are no longer exactly orthogonal to the signal. The effect this can be expected to produce on the main output, is explored here. To simplify the presentation, we will limit this investigation to the first SFR.

What is modified, as compared to figure 3.3, is an additional small

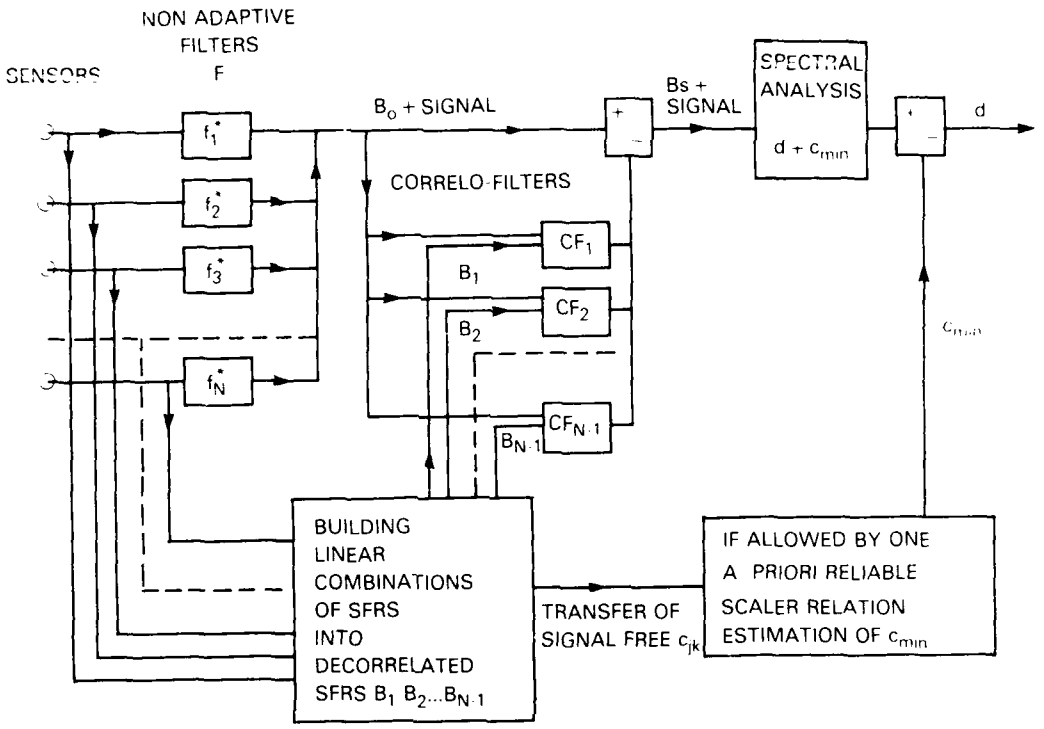


Figure 3.5: General Adaptive Processor with $(N-1)$ Signal-Free References.

amplitude signal superposed on B_1 which can be considered as derived from the main signal through a filter whose frequency response will be denoted by w. Then we have a new csd between the main output and the SFR,

$$\tilde{c}_{01} = c_{01} + dw^*, \tag{3.30}$$

while the spectral density is also modified,

$$\tilde{c}_{11} = c_{11} + d|w|^2. \tag{3.31}$$

Together (3.30) and (3.31) are the quantities estimated by the correlo filter which yields a filter $Q = \tilde{c}_{01}/\tilde{c}_{11}$. The former output noise B_1' is altered to $\tilde{B}_1'$ so that

$$\begin{array}{lll} \mathrm{csd}\left[\tilde{B}_1\tilde{B}_1'\right] & \mathrm{sd}\left[B_1\right]Q & \tilde{c}_{11}Q^*, \quad \text{and} \\ \mathrm{sd}\left[\tilde{B}_1'\right] & \mathrm{sd}\left[\tilde{B}_1\right]Q^2 & \tilde{c}_{11}Q^2. \end{array}$$

On the main output, B_s becomes $\tilde{B}_s$ with:

$$\begin{aligned} \mathrm{sd}\left[\tilde{B}_s\right] &= \mathrm{sd}\left[B_0\right] - \mathrm{csd}\left[B_0\tilde{B}_1'\right] - \mathrm{csd}\left[\tilde{B}_1'B_0\right] + \mathrm{sd}\left[\tilde{B}_1'\right], \\ \mathrm{csd}\left[B_0\tilde{B}_1'\right] &= \mathrm{csd}\left[B_0\tilde{B}_1\right]Q^* = \tilde{c}_{01}Q^*, \\ \mathrm{sd}\left[B_s\right] &= c_{00} - \tilde{c}_{01}Q^* - \tilde{c}_{01}^*Q + \tilde{c}_{11}\left|Q\right|^2. \end{aligned}$$

Hence

$$\mathrm{sd}\left[\tilde{B}_s\right] = c_{00} - \frac{\left|\tilde{c}_{01}\right|^2}{\tilde{c}_{11}},$$

and finally

$$\mathrm{sd}\left[\tilde{B}_s\right] = c_{00} - \frac{\left|c_{01}\right|^2 + d\left(\psi + \psi^*\right) + d^2\left|\psi\right|^2}{c_{11} + d\left|\psi\right|^2}. \tag{3.32}$$

The noise itself is increased by only a second order term when ψ is small,

$$c_{00} - \frac{\left|c_{01}\right|^2}{c_{11} + d\left|\psi\right|^2}.$$

The main point is that some signal now comes through the loop of the correlo-filter. This produces an additive alteration of the spectral-density with the first order component $-d(\psi + \psi^*)/c_{11}$. So the signal spectral density, $d|1 - (v + v^*)/c_{11}|$, may be larger or smaller than the expected density, d, in an unpredictable way since nothing is generally known about the filter v. Small or large, the most spoiling effect of signal leakage into an SFR lies in the fact that the signal spectral density becomes *distorted* by a process which has been conceived to preserve it. That is why other methods have to be considered when the signal wavefront is unknown (chapter 4).

3.11

Nevertheless compromises can be considered in order to make the previous *processing more* robust. Leakage is a consequence of uncertainties about the signal wavefront which can be expressed in terms of subspaces and SFRs. If the signal wavefront cannot be exactly assigned the shape of vector $\mathbf{F}$, at least it will be generally possible to accept it as being located in a limited subspace "around" $\mathbf{F}$. This signal subspace is the generalization of the notion of beamwidth inasmuch as the latter concerns the concept of

aperture limitation and reflects only the uncertainty about the direction of an incoming plane-wave signal. Anyway, the signal subspace has to be determined by a set of vectors $\{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_p\}$, $p = N$, more or less neighboring $\mathbf{F}$.

The signal-free complementary subspace is orthogonal to all these vectors, and is consequently an $(N - p)$-dimensional subspace. With only $(N - p)$ sensor combinations being considered as SFRs the process of removing the unwanted noise from all the p "main-outputs" is likely to be less efficient than before. Indeed the reduction of the SFR number (for a given N) lessens the possibility of exploiting an ideal "quasi-singular" CSDM of unwanted noise and reaching the threshold where only a questionable acoustic "background" noise could not be eliminated. Such is the cost to be paid for the uncertainty, plus the painful repetition of the noise subtraction process from the p main outputs which requires a total of $p(N - p)$ correlo-filters instead of $(N - 1)$.

In short the larger the uncertainty, the larger the signal subspace and the smaller the signal free subspace whose size is a measure of the process efficiency. Total ignorance of vector $\mathbf{F}$ would lead to a generalization of beamforming in an N-dimensional complex space: a mosaic of non overlapping signal subspaces each one being processed after its own complementary one. Let us remember we are considering large arrays of 60 sensors or more. The computational effort seems beyond imagination, a true overdose of processing. The number of sources possessing a wavefront has to be much smaller than N in an array really able to cope with the source and medium complexity. All of these problems might help one to favorably consider chapter 4 in spite of its own perplexing difficulties.

3.12

As a final comment about adaptive processing it must be mentioned that the techniques of adaptive self noise reduction (or canceling) are only a particular case within the framework of the presentation given in this section. As another example consider the vibration of a ship structure in which some of the induced acoustic energy lies within the signal bandwidth and reaches the acoustic sensors of a hull-mounted array. Obviously the wise thing to do first, is to mechanically damp the source of noise. However when the

best possible result has been so obtained residual self noise may persist.

One possible way to try and clear the array of this noise is to use additional sensors (accelerometers) located on or close to the vibrating structures. Such sensors have two very interesting properties for our purposes. They are nearly perfect SFRs because they are mounted inside the hull which makes them practically impervious to interference from acoustic signals produced outside the hull. Also, one can expect that each particular output is highly correlated with that part of the self noise generated by the tested structure.

Therefore a procedure of self noise reduction can be considered. It appears just as an increase in the "size" of the array, each additional sensor delivering a SFR. Self noise SFRs may or may not be correlated either between each other or with the SFRs already extracted from the array. If no reliable information is at hand the diagonalization process has to be applied to *all* the SFRs for them to become uncorrelated. They then feed the subsequent noise subtracting process of sections 3.6 and 3.7.

3.13

As a conclusion to chapter 3 the major points of this type of processing are listed.

1. Dropping unreliable assumptions about unwanted noise leads to adaptive processing with the necessary complexity required to track the variable unwanted noise spatial structure. There is no practical difference between unwanted noise sources gifted with a wavefront, and diffuse noise. So the ambiguity linked to the directivity concept and mentioned in chapter 2, disappears.

2. The processor yielding an unaltered signal in a minimum of unwanted noise can be analyzed by starting with the building of a specific sensor combination to carry the signal and $(N - 1)$ signal-free combinations (SFRs). Then the SFRs, generally correlated, are combined in such a way as to decorrelate them. Further processing is a repetition of "noise subtraction" (noise itself and its spectral density) between the main output and the decorrelated SFRs.

3. The possibility of getting a "perfect" result (no additional noise mixed with the signal) is linked to the aptitude of the array to surmount the complexity of the source-medium system. This aptitude is obviously bounded by the size of the array (its own "complexity"). It can be tested with the set of eigenvalues of the CSDM. The smallest eigenvalue has to be negligible as compared to the largest ones.

4. The signal source-vector is known; *but therein lies the performance* limitation. Any slight swerving of the source vector has to be, when unpredictable, counteracted at the cost of a loss of unwanted noise reduction possibilities. Being more matched to the unwanted noise than classical beamforming, the processor is sensitive to any strict assumption about the signal source vector.

5. It is a vanity to expect, in real conditions, a simple way of coping simultaneously with "imperfections" of a given source vector, and an unknown multiplicity of sources. The transposition of beamforming in a N-dimensional space is fantastically complex for large arrays, and useless since the number of sources is much less than the number of sensors, a definition in some way, of an "efficient" array.

6. The processor composed of correlo-filters described in figures 3.2, 3.3, 3.4 and 3.5, is one possible type of minimum variance distortionless filter. For the sake of physical significance, it has been presented and discussed here in hardware form. Implementation in software is also possible if enough computing capacity is available.

Chapter 4

IDENTIFICATION OF SIGNALS WITH UNKNOWN WAVEFRONTS

4.1

We now come to the main goal of this paper. How can we identify the spectral density of a source with a totally unknown wavefront which must also be determined? In other words we will be trying to find the minimum information needed to restore the spectral density after dropping all a priori knowledge concerning both the signal and noise.

In previous sections the source had been characterized by a known vector involving the notions of spectral density and wavefront. Here such a specific characterization will not be assumed. As a consequence there is no longer any way to distinguish between several sources. This puts us in the position of being able to identify either none, or all sources simultaneously. Under these conditions *the concept of what may be a signal has to be redefined to include every source gifted with an a priori unknown wavefront.* It then becomes difficult to separate those wavefronts in which we have some practical interest from those we wish to ignore. Previously the existence of a wavefront which characterizes "interesting sources" has been implicit in all spatial processing. But since it is now the only assumption retained about a signal, it is worth mentioning that *even this constitutes a priori knowledge*. If the existence of wavefronts comes into question the basis of

all spatial processing is removed, including this whole paper.

Obviously some experiments are necessary to determine how much we can rely on the wavefront notion. The characteristic feature of a source gifted with a "perfect" wavefront is a rank-one CSDM, $\vec{\mathbf{S}}\vec{\mathbf{S}}^{\dagger}$ regardless of the the propagation complexity between the source and the array. The dyadic property of this matrix is the result of so called "perfect spatial coherence" which is characteristic of a single "perfect source."

The question which needs to be tested in operational conditions is whether or not a remote source, which is powerful enough to produce a definitely dominant contribution (+20 dB or more) on the output of *every sensor*, really yields a rank-one experimental CSDM. From such experiments the maximum array size for which the dyadic source approximation will remain acceptable can be determined to some degree of precision. We expect this array size limitation to be much larger than the sort of apertures on which every wavefront has to be a plane or spherical wave. Hopefully we are now in the realm of "large arrays" which has already been mentioned in chapter 3 and will be more precisely presented in section 4.3.

4.2

It is now clear precisely what properties a perfect source signal has. Noise is everything else, a mixture of every kind of diffuse source not possessing the wavefront notion, with a CSDM generally exhibiting a rank which approaches its order. We thus have to disentangle an acoustic field made of an unknown number of "perfect sources" from a "noise" with no characterized wavefront.

We have already noticed in chapter 3 that the *apparent* spectral density, d, and normalized wavefront vector, $\vec{\mathbf{F}}$, of a "perfect source" are the most we can expect to determine using an array. Since (2.7) is $\vec{\mathbf{S}} = d^{\frac{1}{2}}\vec{\mathbf{F}}$, vectors $\vec{\mathbf{S}}$ and $\vec{\mathbf{F}}$ may both be represented as *column* matrices such that $\vec{\mathbf{F}}^{\dagger}\vec{\mathbf{F}} = 1$. We now assume that we have a *limited number of uncorrelated perfect sources* which produce a field received by a large array. By limited we mean the number of sources, P, is less than the number of sensors in the array, $P < N$, which clarifies our notion of a large array.

The CSDM of the perfect sources we shall be working with can then be

represented as a sum of dyads,

$$\vec{\mathbf{C}}_s = \sum_{p=1}^{P} \vec{\mathbf{S}}_p \vec{\mathbf{S}}_p^{\dagger}, \tag{4.1}$$

or equivalently,

$$\vec{\mathbf{C}}_s = \sum_{p=1}^{P} d_p \vec{\mathbf{F}}_p \vec{\mathbf{F}}_p^{\dagger}. \tag{4.2}$$

The rank of this matrix is P, since every vector orthogonal to all the $\vec{\mathbf{F}}_p$ is transformed into a null vector. This statement assumes the set of vectors $\vec{\mathbf{F}}_p$ have no linear dependence which would reduce the rank of $\vec{\mathbf{C}}_s$. This depends somewhat on propagation conditions and can occur if two farfield sources and the array should happen to lie upon a straight line. In practice this is not a stringent assumption. But before exploring matrix $\vec{\mathbf{C}}_s$ we have to extract it from the raw data matrix, the data CSDM $\vec{\mathbf{C}}_r$.

4.3

The data CSDM includes the source dyads plus the CSDM of diffuse noise *supposed to be uncorrelated with any sources.* The net result is that it cannot have a rank lower than its order N. As already mentioned in chapter 3 we can never find an experimental CSDM with an *exactly* null eigenvalue. The only serious accident which can happen to a data CSDM would be the occurrence of a *negative* eigenvalue. This would be unacceptable because it means the correaponding eigenvector is defining a set of filters which passes a *negative* spectral density on its output. Fortunately this accident cannot happen to the maximum likelihood estimate of a data CSDM because such an estimate is the sum of many data dyads. We then have an estimated noise CSDM with the right features required by physics: Hermitian symmetry and real positive eigenvalues.

It is time now to come back to the notion of "large" arrays which involves the concept of an *array possessing enough sensors to surmount the complexity of a system composed of perfect sources plus some diffuse noise structure.* We must emphasize that in a stable situation, *the smallest eigenvalue of a data CSDM generally decreases when the number of sensors increases.* Let us try to illustrate it, with a minimum of mathematical symbolism.

Recalling that eigenvalues represent spectral density, it is well known that the filter combination derived from a normalized vector which *delivers the minimum spectral density* is precisely the combination derived from the normalized eigenvector associated with the minimum eigenvalue.

Indeed if we look for the combination derived from a normalized vector $\vec{\mathbf{v}}$ such that $\vec{\mathbf{v}}^{\dagger}\vec{\mathbf{v}} = 1$, the minimum spectral density of the output $\vec{\mathbf{v}}^{\dagger}\vec{\mathbf{C}}_r\vec{\mathbf{v}}$ is given by a solution to the type of variations problem previously seen in both chapters 2 and 3. The solution is now well known:

$$\begin{aligned} \vec{\mathbf{C}}_r\vec{\mathbf{v}} &= \lambda\vec{\mathbf{v}} \\ \vec{\mathbf{v}}^{\dagger}\vec{\mathbf{C}}_r\vec{\mathbf{v}} &= \lambda \end{aligned} \tag{4.3}$$

with λ a scalar. Necessarily $\vec{\mathbf{v}}$ is an eigenvector and λ is the corresponding eigenvalue. Since we want a minimum output from the filter combination, λ must be the minimum eigenvalue and $\vec{\mathbf{v}}$ the corresponding eigenvector.

In short *the minimum eigenvalue is the minimum output which can be derived from a normalized combination of sensors.* Such a minimization applies to the spectral density at *every* frequency. Let us assume we have built a normalized combination for an N sensor array, with B_0 its output and further assume we can obtain an additional $(N+1)^{\text{th}}$ sensor with output B_1. There is no reason for B_1 to be uncorrelated with B_0. Therefore using a correlo-filter such as the one described in figure 3.3 of chapter 3, with B_0 and B_1 at the same places, we know it is possible to reduce the spectral density of B_0 to a smaller one, B_s. Although the notions of known signal wavefront and SFR have vanished, the CF is not prevented from playing its basic part. We also remark that the correlo-filter effect does not depend at all on the level of B_1, but only on the amount of correlation between B_0 and B_1. The level can be adjusted to a value corresponding to a normalized $(N+1)$-sensor combination because an arbitrary filter is allowed at the SFR input of the correlo-filter.

Thus we have built *a normalized combination of* $(N+1)$ *sensors with an output spectral level smaller than the smallest obtainable with* N *sensors.* And it is not even the smallest possible one with $(N+1)$ sensors, since this is given by the smallest eigenvalue of the $(N+1)$ sensor CSDM. In short, *when the number of sensors increases the smallest eigenvalue decreases*, or at the worst remains stable. Adding more sensors increases the possibility of having a large span of eigenvalues with the smallest one becoming much smaller than the largest one. As a limit it leads either to a singular CSDM

with a vanishing smallest eigenvalue, or to a stabilization of the smallest eigenvalues which are repeated for each additional sensor. In both cases it is a test that the array has become large enough to overcome the complexity of the system of sources and the noise structure.

4.4

Having clarified what we mean by a *large* array, we shall now *assume such arrays for the remainder of this paper.* Since the array must separate sources from the noise and also describe the source wavefronts, it is unlikely that a simple linear array, even a long one, could cope with a fine description of intricate wavefronts. Volumetric arrays are more desirable for our purposes since they can be arranged to break any symmetry which may promote linear dependence between the wavefronts of two distinct sources.

We now have to estimate the number of sources, P, and in the absence of *any* a *priori* information it is clear that we will have to gamble. If a set of several eigenvalues with the same minimum value is supporting the acoustical *background* noise (this is improbable) the difference between the data CSDM and $\epsilon_{\min}\ddot{\mathbf{I}}_N$,

$$\ddot{\mathbf{C}}_r - \epsilon_{\min}\ddot{\mathbf{I}}_N = \ddot{\mathbf{C}}_s \tag{4.4}$$

is a singular matrix of rank "N minus the number of equal eigenvalues $\epsilon_{\min}$." Here $\epsilon_{\min}$ is this minimum eigenvalue and $\ddot{\mathbf{I}}_N$ the unit matrix of order N. It is then possible to consider $\ddot{\mathbf{C}}_s$ as the perfect sources matrix of (4.1) and (4.2). Unfortunately there is little chance the noise structure can be simply described by $\ddot{\mathbf{I}}_N$. Several more elaborate approaches have been suggested [19] [23] which are all smart, but all based on some arbitrary option.

The principle behind the above approach is a parameterized noise CSDM whose general form is known but where several parameters are left free, particularly a scale factor. Such a parameterized noise CSDM may be written as

$$\ddot{\mathbf{M}} = \sigma\ddot{\mathbf{G}}(a, b) \tag{4.5}$$

where σ is a scale factor, $\ddot{\mathbf{G}}$ the general form of a normalized CSDM, with a and b as free specific parameters. Then in the difference matrix

$$\ddot{\mathbf{C}}_s \quad \ddot{\mathbf{C}}_r \quad \ddot{\mathbf{M}}, \tag{4.6}$$

the values of the parameters σ, a and b, are varied to obtain the lowest possible rank of $\ddot{\mathbf{C}}_x$. In this particular case the rank is $N - 3$ but if the general noise matrix $\ddot{\mathbf{G}}$, is really close to the physics of the situation it might well happen that the rank of $\ddot{\mathbf{C}}_x$ will be still lower. We will not comment about the computing difficulties in solving (4.6) since it is already done in the papers by Bienvenu and Kopp [19,20,23,27], as well as by others [26].

If the discovery of a noise matrix best fitted to the real world should prove to be difficult, a generalization could be suggested where $\ddot{\mathbf{G}}$ would be replaced by a sum of weighted noise matrices corresponding to different structures of uncorrelated noise. For example

$$\ddot{\mathbf{C}}_x = \epsilon_{\min}\ddot{\mathbf{I}}_N + \sigma_1\ddot{\mathbf{G}}_1(a_1, b_1) + \sigma_2\ddot{\mathbf{G}}_2(a_2, b_2) \tag{4.7}$$

is one possibility in which all the free parameters could be adjusted to minimize the rank. Again, the success of representing the noise matrix in this way is reflected by the degree of rank reduction in matrix $\ddot{\mathbf{C}}_x$. As we try to increase the number of noise matrices, a natural limitation should appear in the finding of insignificant scale factors σ, a test that we have *exhausted the possibilities* of noise representation.

Anyway the field is wide open to improve such methods or suggest other ones, and experience will probably bring reliable simplifications. For those who immediately need any method, the only resource is to carefully study the eigenvalues of $\ddot{\mathbf{C}}_x$ listed in decreasing order. The smallest ones may be so small that they could be considered negligible. In other cases one can detect a sharp difference between a subset of large eigenvalues and a subset of small ones. P would then be the number of large ones. If need be, nothing prevents us from building a parameterized model of the generalized noise matrix which is dependent on several free parameters, i.e., $\sigma\ddot{\mathbf{G}}(a, b, \ldots)$. What has been suggested requires *minimal assumptions* because only the general form of the noise spatial structure has to be postulated. Nevertheless, as mentioned in section 4.1, these assumptions have to be obtained from total ignorance using a priori knowledge.

4.5

The procedure for determining the matrix, $\ddot{\mathbf{C}}_x$, of the fully coherent spatial sources requires a parametric model to eliminate the second order noise

statistics. Any further processing of matrix $\bar{\bar{\mathbf{C}}}_s$ can now be qualified as "high-resolution" with regard to the "perfect" sources. Matrix $\bar{\bar{\mathbf{C}}}_s$ is bound to the source vectors $\bar{\mathbf{S}}_p$ by (4.1). The problem is to determine whether or not we can explicitly find the vectors $\bar{\mathbf{S}}_p$ when $\bar{\bar{\mathbf{C}}}_s$ is known. See references [8,24].

Since the order of $\bar{\bar{\mathbf{C}}}_s$ is N and its rank is P, there are P nonzero eigenvalues and the corresponding eigenvectors describe a P-dimensional subspace of the N-dimensional space. We will now show the wavefront vectors, $\bar{\mathbf{S}}_p$, also span the same P-dimensional subspace. In the $(N - P)$-dimensional complementary subspace the eigenvalues are all nulls so the associated eigenvectors are indeterminate which we shall find to be unimportant.

The set of nonzero eigenvalues, in decreasing order, is:

$$\epsilon_1, \epsilon_2, \ldots \epsilon_p, \ldots \epsilon_P$$

and the corresponding normalized eigenvectors (each a column matrix with N components, like the $\bar{\mathbf{S}}_p$) are:

$$\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2, \ldots \bar{\mathbf{u}}_p, \ldots \bar{\mathbf{u}}_P$$

with the usual orthonormal conditions:

$$\begin{aligned} \bar{\mathbf{u}}_p^\dagger \bar{\mathbf{u}}_p &= 1; \\ \bar{\mathbf{u}}_p^\dagger \bar{\mathbf{u}}_q &= 0, \quad p \neq q. \end{aligned}$$

Then matrix $\bar{\bar{\mathbf{C}}}_s$ may be written in canonical form as justified in appendix E.

$$\bar{\bar{\mathbf{C}}}_s = \sum_{p=1}^{P} \epsilon_p \left[\bar{\mathbf{u}}_p \bar{\mathbf{u}}_p^\dagger \right], \tag{4.8}$$

where the $\left[\bar{\mathbf{u}}_p \bar{\mathbf{u}}_p^\dagger \right]$ are dyadics of the vector $\bar{\mathbf{u}}_p$. Comparing with (4.1) we can write,

$$\sum_{p=1}^{P} \bar{\mathbf{S}}_p \bar{\mathbf{S}}_p^\dagger = \sum_{p=1}^{P} \epsilon_p \left[\bar{\mathbf{u}}_p \bar{\mathbf{u}}_p^\dagger \right]; \tag{4.9}$$

both sides are the same matrix. If we now look at the transform of vector $\bar{\mathbf{u}}_1$, we have

$$\sum_{p=1}^{P} \bar{\mathbf{S}}_p \left(\bar{\mathbf{S}}_p^\dagger \bar{\mathbf{u}}_1 \right) = \sum_{p=1}^{P} \epsilon_p \bar{\mathbf{u}}_p \left(\bar{\mathbf{u}}_p^\dagger \bar{\mathbf{u}}_1 \right). \tag{4.10}$$

where scalars are shown between parenthesis. Because of the above orthonormality conditions between these $\vec{\mathbf{u}}_p$, the right side of (4.10) is reduced to $(c_1)\vec{\mathbf{u}}_1$. Therefore vector $\vec{\mathbf{u}}_1$ appears as a linear combination of the vectors $\vec{\mathbf{S}}_p$. As a consequence $\vec{\mathbf{u}}_1$ *necessarily belongs to the subspace spanned by* $\vec{\mathbf{S}}_p$. The same could be shown for every $\vec{\mathbf{u}}_p$. Conversely the unknown $\vec{\mathbf{S}}_p$ *belong to the P-dimensional subspace described by the known* $\vec{\mathbf{u}}_p$. This is the first important feature of the vectors $\vec{\mathbf{S}}_p$; it is not the only one. In this subspace the $\vec{\mathbf{S}}_p$ are bound by several other relations to be discussed next. Incidently, the above discussion disproves a possible misconception that individual sources can be described by individual eigenvectors.

4.6

Things will be easier if we accept a preliminary "theorem" which is demonstrated in appendix F. Let us consider a set of P vectors $\vec{\mathbf{V}}_p$ describing a P-dimensional subspace within an N dimensional vector space, and also a diagonal matrix $\vec{\vec{\Pi}}$ *of order* N, built as follows:

$$\vec{\vec{\Pi}} \quad \left(\begin{array}{cccc} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ & & & 0 \\ & 0 & & \ddots \\ & & & & 0 \end{array} \right) \begin{array}{l} \left. \begin{array}{l} \\ \\ \end{array} \right\} P \text{ rows} \\ \left. \begin{array}{l} \\ \\ \end{array} \right\} N - P \text{ rows} \end{array} \tag{4.11}$$

The matrix $\vec{\vec{\Pi}}$ may now be written as a sum of dyads,

$$\vec{\vec{\Pi}} = \sum_{p=1}^{P} \vec{\mathbf{V}}_p \vec{\mathbf{V}}_p^{\dagger}, \tag{4.12}$$

where vectors $\vec{\mathbf{V}}_p$ have the following features:

1. Only the first P components of each $\vec{\mathbf{V}}_p$ differ from zero. In other words vectors $\vec{\mathbf{V}}_p$ describe a P-dimensional subspace spanned by the first P axes of the full N-dimensional space.

2. The $\vec{\mathbf{V}}_p$ will be chosen to satisfy orthonormality conditions, $\vec{\mathbf{V}}_p^{\dagger}\vec{\mathbf{V}}_p = 1$; $\vec{\mathbf{V}}_p^{\dagger}\vec{\mathbf{V}}_q = 0$, $p \neq q$.

This being accepted, we know that the matrix $\ddot{\mathbf{C}}_s$ can be written as

$$\ddot{\mathbf{C}}_s = \ddot{\mathbf{U}}\ddot{\mathbf{E}}\ddot{\mathbf{U}}^\dagger \tag{4.13}$$

(see appendix C), where $\ddot{\mathbf{E}}$ is the diagonal matrix of eigenvalues, ϵ_p, in decreasing order with the last $(N - P)$ ones being null, and $\ddot{\mathbf{U}}$ is the N-dimensional unitary matrix whose *columns* are the normalized eigenvectors of $\ddot{\mathbf{C}}_s$, set out in the same order as the corresponding eigenvalues in $\ddot{\mathbf{E}}$. Although the last $(N - P)$ columns are indeterminate they are arbitrary eigenvectors and obviously $\ddot{\mathbf{U}}^\dagger\ddot{\mathbf{U}} = \ddot{\mathbf{I}}_N$ a relation specific to unitary matrices. ($\ddot{\mathbf{I}}_N$ is the N-dimensional identity matrix.)

Since the matrix $\ddot{\mathbf{E}}$ is not invertible let us consider a diagonal matrix of real numbers, $\ddot{\mathbf{D}} = \ddot{\mathbf{D}}^T$, obtained from $\ddot{\mathbf{E}}$ by replacing the $(N - P)$ diagonal zeros of $\ddot{\mathbf{E}}$ with some arbitrary scalar, here chosen to be 1. It is clear that we can write,

$$\ddot{\mathbf{E}} = \ddot{\mathbf{D}}^{\frac{1}{2}}\ddot{\Pi}\ddot{\mathbf{D}}^{\frac{1}{2}} \tag{4.14}$$

where $\ddot{\mathbf{D}}^{\frac{1}{2}}$ is partly composed of the eigenvalue square roots. Hence, following (4.13),

$$\ddot{\mathbf{C}}_s = \ddot{\mathbf{U}}\ddot{\mathbf{D}}^{\frac{1}{2}}\ddot{\Pi}\ddot{\mathbf{D}}^{\frac{1}{2}}\ddot{\mathbf{U}}^\dagger, \tag{4.15}$$

which may be solved for $\ddot{\Pi}$ as follows:

$$\begin{aligned} \ddot{\mathbf{U}}^\dagger\ddot{\mathbf{C}}_s\ddot{\mathbf{U}} &= \ddot{\mathbf{D}}^{\frac{1}{2}}\ddot{\Pi}\ddot{\mathbf{D}}^{\frac{1}{2}}, \\ \ddot{\mathbf{D}}^{-\frac{1}{2}}\ddot{\mathbf{U}}^\dagger\ddot{\mathbf{C}}_s\ddot{\mathbf{U}}\ddot{\mathbf{D}}^{-\frac{1}{2}} &= \ddot{\Pi}. \end{aligned} \tag{4.16}$$

But we know that,

$$\ddot{\mathbf{C}}_s = \sum_{p=1}^{P} \vec{\mathbf{S}}_p\vec{\mathbf{S}}_p^\dagger, \quad \text{from (4.1).}$$

Then since $\ddot{\mathbf{D}}, \ddot{\mathbf{U}}$ and $\ddot{\Pi}$ are common for all $\vec{\mathbf{S}}_p$, we have

$$\ddot{\Pi} \quad \sum_{p=1}^{P} \left[\ddot{\mathbf{D}}^{-\frac{1}{2}}\ddot{\mathbf{U}}^\dagger\vec{\mathbf{S}}_p\right]\left[\vec{\mathbf{S}}_p^\dagger\ddot{\mathbf{U}}\ddot{\mathbf{D}}^{-\frac{1}{2}}\right]. \tag{4.17}$$

Let

$$\vec{\mathbf{V}}_p \quad \ddot{\mathbf{D}}^{-\frac{1}{2}}\ddot{\mathbf{U}}^\dagger\vec{\mathbf{S}}_p. \tag{4.18}$$

then we have

$$\ddot{\Pi} = \sum_{p=1}^{P} \vec{\mathbf{V}}_p \vec{\mathbf{V}}_p^{\dagger}. \tag{4.19}$$

Therefore the $\vec{\mathbf{V}}_p$ have all the features listed in items 1 and 2 at the beginning of this section. We can then derive features of the $\vec{\mathbf{S}}_p$ by solving (4.18). Since $\ddot{\mathbf{U}}^{\dagger} = \ddot{\mathbf{U}}^{-1}$ (4.18) yields

$$\vec{\mathbf{S}}_p = \ddot{\mathbf{U}} \ddot{\mathbf{D}}^{\frac{1}{2}} \vec{\mathbf{V}}_p, \tag{4.20}$$

which expresses *every* $\vec{\mathbf{S}}_p$ as a function of $\ddot{\mathbf{C}}_s$ and of an N-dimensional vector $\vec{\mathbf{V}}_p$ gifted with the features discussed above in items 1 and 2.

Since only the first P components of $\vec{\mathbf{V}}_p$ differ from zero, the same must be true for vector $\vec{\mathbf{V}}'_p = \ddot{\mathbf{D}}^{\frac{1}{2}} \vec{\mathbf{V}}_p$. The diagonal elements of $\ddot{\mathbf{D}}^{\frac{1}{2}}$ from the $(P+1)^{\text{th}}$ to the N^{th}, those which had been arbitrarily chosen, in fact disappear. Matrix $\ddot{\mathbf{U}}$ is made of column vectors which are the normalized eigenvectors of $\ddot{\mathbf{C}}_s$. The indeterminate eigenvectors are neutralized because when multiplying $\ddot{\mathbf{U}}$ by the vector $\vec{\mathbf{V}}'_p$, the elements of $\ddot{\mathbf{U}}$ from the $(P+1)^{\text{th}}$ to the N^{th} column are always multiplied by the zero elements of vector $\vec{\mathbf{V}}'_p$.

Equation (4.20) may be reduced to its essential components,

$$\vec{\mathbf{S}}_p = \ddot{\mathbf{U}}^P \ddot{\Delta}^P \vec{\mathbf{X}}_p, \tag{4.21}$$

where:

$\ddot{\mathbf{U}}^P$ is the rectangular matrix (N rows $\times$ P columns) whose columns are the first P normalized eigenvectors of $\ddot{\mathbf{C}}_s$.

$\ddot{\Delta}^P$ is a square (P, P) matrix whose diagonal elements $\{c_1^{\frac{1}{2}} \ldots c_P^{\frac{1}{2}}\}$, are the square roots of the nonzero eigenvalues of C_s, arranged in decreasing order and corresponding to the eigenvectors of $\ddot{\mathbf{U}}$.

$\vec{\mathbf{X}}_p$ is a P-component vector (column matrix) whose components are the first P components of $\vec{\mathbf{V}}_p$. The set of $\vec{\mathbf{X}}_p$ is gifted with the same features listed in items 1 and 2 as the set of $\vec{\mathbf{V}}_p$.

Equation (4.21) is coherent in the sense that both the left and right sides should be column matrices with N rows i.e. dimension $(N \times 1)$. Indeed:

$$\begin{aligned} \ddot{\mathbf{U}}^P &\text{ is } (N \times P), \\ \ddot{\Delta}^P &\text{ is } (P \times P), \\ \vec{\mathbf{X}}_p &\text{ is } (P \times 1); \end{aligned}$$

so the final result is $(N \times 1)$ as expected.

There is a bit more to say about (4.21). When computing the eigenvectors of a CSDM, we can get each of them only up to a phase factor $\exp(i\varphi_p)$. If we choose one possible determination of $\vec{\mathbf{U}}^P$, denoted by $\vec{\mathbf{U}}_0^P$, the most general expression of $\vec{\mathbf{U}}^P$ will be

$$\vec{\mathbf{U}}^P = \vec{\mathbf{U}}_0^P \vec{\Phi}^P \tag{4.22}$$

here $\vec{\Phi}^P$ is the diagonal $(P \times P)$ matrix of the phase factors. Particularly $(\vec{\Phi}^P)^\dagger \vec{\Phi}^P = \vec{\mathbf{I}}_P$, where $\vec{\mathbf{I}}_P$ is the identity matrix in a P-dimensional space. This gives (4.21) the form $\vec{\mathbf{S}}_p = \vec{\mathbf{U}}_0^P \vec{\Phi}^P \vec{\Delta}^P \vec{\mathbf{X}}_p$, or permuting the two diagonal matrices,

$$\vec{\mathbf{S}}_p = \vec{\mathbf{U}}_0^P \vec{\Delta}^P \vec{\Phi}^P \vec{\mathbf{X}}_p. \tag{4.23}$$

Now the specific features of $\vec{\mathbf{X}}_p$ necessarily react on the vector $\vec{\mathbf{Z}}_p \equiv \vec{\Phi}^P \vec{\mathbf{X}}_p$.

$$\text{Since} \quad \vec{\mathbf{X}}_p^\dagger \vec{\mathbf{X}}_p = 1, \qquad \text{then} \quad \vec{\mathbf{Z}}_p^\dagger \vec{\mathbf{Z}}_p = \vec{\mathbf{X}}_p^\dagger \left(\left(\vec{\Phi}^P \right)^\dagger \vec{\Phi}^P \right) \vec{\mathbf{X}}_p = 1.$$

$$\text{Since} \quad \vec{\mathbf{X}}_p^\dagger \vec{\mathbf{X}}_q = 0, \quad p = q, \quad \text{then} \quad \vec{\mathbf{Z}}_p^\dagger \vec{\mathbf{Z}}_q = \vec{\mathbf{X}}_p^\dagger \left(\left(\vec{\Phi}^P \right)^\dagger \vec{\Phi}^P \right) \vec{\mathbf{X}}_q = 0.$$

These important features have been conveyed from vectors $\vec{\mathbf{V}}_p$ to vectors $\vec{\mathbf{Z}}_p$ through vectors $\vec{\mathbf{X}}_p$. Therefore, $\vec{\mathbf{Z}}_p$ *is the* p^{th}*-column of a* $(P \times P)$ *unitary matrix in the* P*-dimensional subspace of the first* P *eigenvectors of the CSDM.*

Ultimately

$$\boxed{\vec{\mathbf{S}}_p = \vec{\mathbf{U}}_0^P \vec{\Delta}^P \vec{\mathbf{Z}}_p}, \tag{4.24}$$

which expresses every $\vec{\mathbf{S}}_p$ as a function of $\vec{\mathbf{C}}_s$ and a partly arbitrary vector, gifted with the features mentioned above in items 1 and 2. The fundamental equation (4.24) can be visualized as below:

$$\underset{(N \times 1)}{\begin{pmatrix} \vdots \\ \vdots \\ \vec{\mathbf{S}}_p \\ \vdots \\ \vdots \end{pmatrix}} \quad \underset{(N \times P)}{\begin{pmatrix} \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \vec{\mathbf{u}}_1 & \cdots & \vec{\mathbf{u}}_p & \cdots & \vec{\mathbf{u}}_P \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \end{pmatrix}} \quad \underset{(P \times P)}{\begin{pmatrix} \epsilon_1^{\frac{1}{2}} & & & & \\ & \ddots & & 0 & \\ & & \epsilon_p^{\frac{1}{2}} & & \\ & 0 & & \ddots & \\ & & & & \epsilon_P^{\frac{1}{2}} \end{pmatrix}} \quad \underset{(P \times 1)}{\begin{pmatrix} \vdots \\ \vdots \\ \vec{\mathbf{Z}}_p \\ \vdots \\ \vdots \end{pmatrix}}$$

Let us have:

$$\begin{array}{llll} u_k^n & \text{the } n^{\text{th}} & \text{component of } \vec{\mathbf{u}}_k, & \text{with } 1 \le n \le N, \\ z_{pk} & \text{the } k^{\text{th}} & \text{component of } \vec{\mathbf{Z}}_p, & \text{with } 1 \le k \le P, \\ s_{pn} & \text{the } n^{\text{th}} & \text{component of } \vec{\mathbf{S}}_p, & \\ \sigma_k & = & \epsilon_k^{\frac{1}{2}}. & \end{array}$$

Then (4.24) becomes

$$\boxed{s_{pn} = \sum_{k=1}^{P} u_k^n \sigma_k z_{pk}}. \tag{4.25}$$

One can remark that the n^{th} component of the $\vec{\mathbf{S}}_p$ depends on all the components of vector $\vec{\mathbf{Z}}_p$. Therefore any s_{pn} depends on all arbitrary scalars which could be found in vector $\vec{\mathbf{Z}}_p$ when building the arbitrary unitary matrix $\ddot{\mathbf{Z}}$. Furthermore, s_{pn} also depends on *all* the eigenvalues.

4.7

The first conclusion to be drawn from (4.24) is that an exhaustive use of the CSDM *is not enough* to determine the P source vectors. What is lacking is a $(P \times P)$ *unitary* matrix $\ddot{\mathbf{Z}}$ whose columns are the partly arbitrary vectors $\vec{\mathbf{Z}}_p$ which are necessary to determine $\vec{\mathbf{S}}_p$. Using a priori information we have to choose a matrix $\ddot{\mathbf{Z}}$, such that $\ddot{\mathbf{Z}}^\dagger \ddot{\mathbf{Z}} = \ddot{\mathbf{I}}_P$.

One may wonder why a unitary matrix is needed. The physics of this point is illustrated by figure B.1 in appendix B. As mentioned there, if the matrix of filters, $\ddot{\mathbf{G}}$, is given by the normalized eigenvectors of $\ddot{\mathbf{C}}_s$, the P nonzero outputs (the zero ones yield nulls) are *not correlated* and their spectral densities are respectively the known eigenvalues. On the other hand we remember that the P sources are postulated to be independent. That could lead one to think that we have at last disentangled the sources from each other. Unfortunately this is not true in general. A pair of uncorrelated outputs can be thought of as two mixtures, each with contributions from all of the sources, in which partial correlations coming from of one source occur in both mixtures. But it happens that these partial correlations exactly cancel each other. Therefore each eigenvalue represents a combination of all P source spectral densities.

Now since the eigenvalues are known, there is a posibility to go a bit further by whitening all the P outputs, thus flattening the spectrum to a uniform density. This is the purpose of matrix $\vec{\Delta}^P$ in (4.24). Thus, the most we can obtain using only the CSDM, are P spectral-normalized mixtures of the sources. That explains how the separation of sources requires assumptions which can be expressed as a unitary or "rotation" matrix.

One case where a solution can be reached without additional assumptions occurs when $P = 1$, where obviously the wavefront is the eigenvector $\vec{u}_1$, the spectral density ϵ_1, and the only source $\vec{\mathbf{S}}_1$. The matrix $\vec{\mathbf{Z}}$ is then reduced to its first element, whose modulus is equal to 1, and therefore corresponds to an arbitrary phase factor. This is of no consequence since a wavefront represents the relative phases of its components and is not changed by a phase facor common to all the elements.

Another bold case where a solution appears from (4.24) occurs when the array is so large and super-resolving that all the source-vectors are virtually orthogonal. Then the eigenvectors are the wavefronts and the corresponding eigenvalues are the spectral densities. Under these conditions the matrix $\vec{\mathbf{Z}}$ is reduced to a diagonal matrix of arbitrary phase factors, one per source vector.

Before discussing the general case it might be interesting to check that (4.21) is compatible with (4.13) when the useless components of the $(N-P)$-dimensional complementary subspace are removed from matrix $\vec{\mathbf{C}}_s$. If we express $\vec{\mathbf{S}}_p\vec{\mathbf{S}}_p^\dagger$ in the same form as (4.24) we get

$$\vec{\mathbf{S}}_p\vec{\mathbf{S}}_p^\dagger = \vec{\mathbf{U}}_0^P\vec{\Delta}^P\vec{\mathbf{Z}}_p\vec{\mathbf{Z}}_p^\dagger\vec{\Delta}^P\left(\vec{\mathbf{U}}_0^P\right)^\dagger. \tag{4.26}$$

The matrix $\vec{\mathbf{Z}}_p\vec{\mathbf{Z}}_p^\dagger$ is the dyadic of $\vec{\mathbf{Z}}_p$. Now

$$\vec{\mathbf{C}}_s = \sum_{p=1}^{P}\vec{\mathbf{S}}_p\vec{\mathbf{S}}_p^\dagger = \vec{\mathbf{U}}_0^P\vec{\Delta}^P\left[\sum_{p=1}^{P}\vec{\mathbf{Z}}_p\vec{\mathbf{Z}}_p^\dagger\right]\vec{\Delta}^P\left(\vec{\mathbf{U}}_0^P\right)^\dagger. \tag{4.27}$$

The matrix between brackets is the *identity* matrix of order P. Indeed any

column vector $\vec{\mathbf{v}}$ is transformed as

$$\left[\sum_{p=1}^{P} \vec{\mathbf{Z}}_p \vec{\mathbf{Z}}_p^{\dagger}\right] \vec{\mathbf{v}} = \sum_{p=1}^{P} \vec{\mathbf{Z}}_p \left(\vec{\mathbf{Z}}_p^{\dagger} \vec{\mathbf{v}}\right). \tag{4.28}$$

The scalar $(\vec{\mathbf{Z}}_p^{\dagger} \vec{\mathbf{v}})$ is precisely the "projection" of $\vec{\mathbf{v}}$ on the unit vector of the p^{th} axis. So (4.28) simply means the transform of $\vec{\mathbf{v}}$ is the vector sum of its own components along the P axis, and therefore $\vec{\mathbf{v}}$ *itself*. Hence

$$\vec{\mathbf{C}}_s = \vec{\mathbf{U}}_0^P \left[\vec{\Delta}^P\right]^2 \left[\vec{\mathbf{U}}_0^P\right]^{\dagger}, \tag{4.29}$$

where $[\vec{\Delta}^P]^2$ is the diagonal matrix of the eigenvalues $\vec{\mathbf{E}}$ in the only subspace where they are not nulls. $\vec{\mathbf{U}}_0^P$ is the unitary matrix $\vec{\mathbf{U}}$ reduced to the useful eigenvectors. Now we note the form of (4.13) and (4.29) are very much the same.

4.8

The problem is now to correctly assess the degrees of freedom required by the "arbitrary" unitary matrix $\vec{\mathbf{Z}}$ in order to determine the necessary *a priori assumptions*. While enough assumptions are needed to compute the $\vec{\mathbf{S}}_p$ using the full power of the array, we wish to avoid overconstraining the problem to preserve the information contained in the data CSDM. To do this a unitary matrix ($P \times P$) is built of P^2 complex scalars which are equivalent to $2P^2$ real scalars. The constraints between these scalars are:

1. Every column must be a normalized vector giving one real scalar relation per $\vec{\mathbf{S}}_p$ for a total of P relations.

2. Every column must be orthogonal to all the others, which corresponds to $P(P-1)/2$ relations between complex numbers and therefore a total of $P(P-1)$ relations between real scalars.

So the $2P^2$ real scalars of $\vec{\mathbf{Z}}$ are bound by P^2 relations, and the arbitrary freedom within matrix $\vec{\mathbf{Z}}$ would seem to be dependent on P^2 *real* scalars. However we have seen that the phase-factor of each column $\vec{\mathbf{Z}}_p$ of $\vec{\mathbf{Z}}$ plays no role in our analysis thus allowing *one* real scalar to remain undetermined.

As a result the number of real scalars needed to get all the $\vec{\mathbf{S}}_p$ is $P^2 - P = P(P-1)$.

If we look at the set of P vectors given by (4.24), each corresponds to N complex scalar equations, or equivalently to $2N$ real scalar equations. For the whole set we have, $2NP$ real scalar equations. What is the number of the unknown parameters? We already know there are $P(P-1)$ real scalars introduced by the vectors $\vec{\mathbf{Z}}_p$, gathered in matrix $\ddot{\mathbf{Z}}$. On the left side, all $2NP$ scalars forming the components of the $\vec{\mathbf{S}}_p$ are unknown.

We then have more unknown parameters than equations. That is why *total ignorance* about the $\vec{\mathbf{S}}_p$ prevents any solution to emerge, even with the most extensive use of the CSDM. *Something has to be assumed about the* $\vec{\mathbf{S}}_p$; but any assumption should provide precisely $P(P-1)$ *real scalar equations between the components of the* $\vec{\mathbf{S}}_p$. More assumptions would lead us to impose specific features (relations between the elements) on both the CSDM and the related matrix $\ddot{\mathbf{U}}_0^P \ddot{\Delta}^P$, therefore causing conflicts between the experimental results and the assumptions.

Before trying a general approach using the minimum necessary assumptions, it is interesting to illustrate the previous development with the case $P = 2$. A unitary matrix of order 2 is made originally of 8 real scalars. But the constraints $\ddot{\mathbf{Z}}^\dagger \ddot{\mathbf{Z}} = \ddot{\mathbf{I}}_2$ lead to the classical writing,

$$\ddot{\mathbf{Z}} = \begin{pmatrix} \cos(\psi) & \sin(\psi)\exp(-i\theta) \\ \sin(\psi)\exp(i\theta) & \cos(\psi) \end{pmatrix} \begin{pmatrix} \exp(i\phi_1) & 0 \\ 0 & \exp(i\phi_2) \end{pmatrix}$$

where only 4 real scalars remain. The diagonal matrix on the right side corresponds to the phase factors of the first and second columns. They can remain arbitrary so that only 2 real scalars ψ and θ have to be introduced in the vector equations.

It must be observed that both scalars are present in every element of the matrix and particularly in all columns. The generalization of this for larger values of P is a harbinger of computing difficulties which arise when determining the source vectors as functions of the arbitrary scalars. In the general case we are confronted with finding the number of scalar relations between the components of the $\vec{\mathbf{S}}_p$ which are necessary to complete the set of equations (4.24) with exactly as many unknown parameters as there are available equations. Not violating this condition with reliable, compatible assumptions is hopefully a way to solve the problem of finding source vectors *using the full array power.*

For the case $P = 2$, (4.24) becomes:

$$\begin{aligned} \vec{\mathbf{S}}_1 &= \vec{\mathbf{U}}_0^P \begin{pmatrix} \epsilon_1^{\frac{1}{2}} & 0 \\ 0 & \epsilon_2^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \cos(\psi) \\ \sin(\psi)\exp(i\theta) \end{pmatrix} \\ \vec{\mathbf{S}}_2 &= \vec{\mathbf{U}}_0^P \begin{pmatrix} \epsilon_1^{\frac{1}{2}} & 0 \\ 0 & \epsilon_2^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} -\sin(\psi)\exp(-i\theta) \\ \cos(\psi) \end{pmatrix} \end{aligned} \tag{4.36}$$

These two vector equations between the components of $\vec{\mathbf{S}}_1$ and $\vec{\mathbf{S}}_2$, yield $2N$ complex scalar equations or $4N$ real scalar equations. On the left side of (4.30) the unknown parameters to be computed are precisely the $4N$ real scalars of the components. On the right side, the unknowns (unavoidable whether interesting or not) are ψ and θ. There are two more unknowns than relations so we must either:

1. Arbitrarily forget about ψ and θ by deciding $\psi = 0$ and assume the two source vectors are orthogonal (very large arrays), a rather crude procedure.

2. "Invent" 2 real scalar relations between the $2N$ real scalars of the source vectors. This is not very stringent and is exactly what is required to reach a solution. More assumptions would be too much and would require compatibility with the experimental CSDM.

4.9

Let us first remember that we are considering arrays of some 60 sensors or more, and P up to 20 or 30. Clearly, all the assumptions made in previous sections about the source vectors are of the "overinformation" type, as compared to the minimum sufficient assumptions. For example the "plane-wave assumption" of chapter 2 corresponds to a source vector structure with a complex amplitude and two angles for the direction. A total of 4 real parameters are needed for this vector description. It actually implies $2N - 4$ implicit relations between the $2N$ real scalars of the components giving a total of $P(2N - 4)$ relations for all the source vectors which exceeds the $P(P - 1)$ strictly necessary relations compatible with the experimental CSDM $\vec{\mathbf{C}}_x$. Sometimes the minimum relations are practically anticipated. As an example, for a linear array of equispaced identical sensors, the source

vector CSDM is of the Toeplitz type.[1] The Toeplitz criterion helps separate the source CSDM $\ddot{\mathbf{C}}_s$ from the data CSDM $\ddot{\mathbf{C}}_r$, which contains unwanted noise, thus giving an estimate of P (the problem is to find the largest possible Toeplitz matrix contained in $\ddot{\mathbf{C}}_r$).

Vezzosi [25] has given a full solution under less stringent assumptions. For a source vector, only the amplitude of the components is supposed to be constant. This common-value amplitude and the phases must then be determined. The assumption of a constant amplitude for N components corresponds to $N - 1$ real scalar relations for each source vector and $P(N-1)$ relations for all sources. Although this is a big step forward, it is still overconstrained since it exceeds the minimum $P(P-1)$ relations. Again, the source CSDM is somewhat specific thus giving a clue for the estimation of P.

Other examples of this kind could be proposed more or less based on intuition about specific situations. True, it is painful to refrain from simplifying an enormously difficult problem by adjusting the CSDM to a given shape, and then further refrain from blaming any discrepancies with the experimental CSDM upon a lack of spectral and cross-spectral density estimation accuracy. CSDM shape adjustment becomes even more tempting because it yields an estimate of P, and we must admit that in section 1.1 we could only derive such estimates from conjectures about the set of eigenvalues. But this is only transferring conjectures from one place to another. Clearly compromises may be successful in some situations. Still, adjusting a problem to its solutions leaves us uneasy. Anyway, for better or worse we have chosen to sail, so we shall now try to find a minimum set of assumptions which balance the equations with the unknowns.

4.10

Our goal is now to suggest something about the source vectors or the wavefronts themselves (and not about a model of the medium since we do not pretend to localize). It should be a "minimum" assumption which is both complete and reliable thus allowing us to use the full power of the array as

[1] A constraint imposed on the $\ddot{\mathbf{C}}_s$ matrix by the "plane-wave assumption" made for any N sensor array, has some internal order such as the Toeplitz matrix, only more difficult to express.

expressed in the CSDM through the set of vector equations (4.24).

The assumption proposed here is: *a source vector, which is not necessarily a plane wave, can be described as a coherent sum of plane waves across the aperture of the array.* Such a superposition should approximate the source vector more closely when the number of plane waves is large, but this number is limited by the number of sensors in the array, N. This assumption may be disputed, but it seems to deserve the qualification "minimum" because as we shall see later:

1. The number of plane waves is actully large in most of the situations we can consider.

2. The geometrical significance of a plane wave is not here an *absolute* preliminary requirement but is only to be a posteriori controlled.

The last remark gives the wavefront model large flexibility. If an expected "plane" wave is found to be non-plane (we shall see how), it still contributes to a parametric model of the wavefront. If the plane-wave feature is confirmed, it gives a clue for localization. That is why we shall name the model a coherent sum of "conditional plane waves," CPW.

Thinking of large fixed arrays, let us now restrict ourselves to an array with N point-like sensors (as compared to the smallest wavelength involved) which have a flat frequency response over the useful bandwidth. Since the array geometry is completely known we can define a reference system in the real three-dimensional space with a point origin and three orthogonal axes. The Cartesian position of the n^{th} sensor is then given by a *known* vector which specifies the length and direction from the origin to the sensor, $\vec{r}_n$.

On the other hand an incoming plane wave is completely described by a normalized 3-dimensional vector $\vec{\alpha}$, $|\vec{\alpha}| = 1$, pointed from the origin *towards* the plane wave (opposite to its propagation direction) representing *two* angles (bearing and tilt). This vector *does not* depend on frequency and is unknown in our problem. A "complex amplitude" A representing the amplitude and phase at the origin-point is also unknown and frequency dependent.

Very classically the phase factor corresponding to the travel time delay from the origin to the sensor, is $\exp[2\pi\imath(\vec{\omega}_n \cdot \vec{\alpha})]$ where the dot "$\cdot$" stands for a scalar inner-product and where the vector $\vec{\omega}_n$ is the vector $\vec{r}_n$ measured in terms of wavelengths. The "length" of vector $\vec{\omega}_n$ is proportional

to frequency but remains completely known in every frequency bin. As we have done throughout this paper, we let the frequency dependence be implicit. Therefore $\vec{\omega}_n$ is a known vector in the 3-dimensional space, while for a "true" plane wave, $\vec{\alpha}$ is unknown, frequency independent, and represents two real scalars. For a CPW, vector $\vec{\alpha}$ will be computed in every frequency bin. Only a posteriori will the CPW be verified as truly being a frequency independent plane wave.

As for the complex amplitude A, obviously $|A|$ is the same on all sensors and at the origin. If we were to represent only one plane wave, A could be a real positive scalar, and the above phase factor would express the phases, relative to that of the origin. But since we shall deal with a sum of *coherent* plane waves, we have to describe the relative phases of several plane waves at the origin, and A has to be a complex, frequency dependent scalar. So then A depends on 2 real scalars, amplitude and phase, both unknown in every frequency bin. The expression for a plane wave on the n^{th} sensor is then

$$A\exp\left[2\pi i\left(\vec{\omega}_n\cdot\vec{\alpha}\right)\right]$$

and depends on 4 real scalars. While two of them *should not* be frequency dependent, this may not always occur.

4.11

Any source vector $\vec{\mathbf{S}}$ described as a sum of K_s coherent plane waves, yields on the n^{th} sensor a component,

$$\sum_{k=1}^{K_s} A_k\exp\left[2\pi i\left(\vec{\omega}_n\cdot\vec{\alpha}_k\right)\right].$$

For some particular source vector $\vec{\mathbf{S}}_p$ with a specific number of CPW, K_p, the n^{th} component is given by

$$s_{pn}=\sum_{k=1}^{K_p} A_{kp}\exp\left[2\pi i\left(\vec{\omega}_n\cdot\vec{\alpha}_{kp}\right)\right]. \tag{4.31}$$

K_p may not be the same for all $\vec{\mathbf{S}}_p$ because each source may have a different multipath arrival structure. This allows us flexibility in the way we express

any loose a priori idea we might have about the relative complexity of the source vectors. The component s_{pn} depends on $4K_p$ real scalars. When we follow (4.31) from one sensor to another only the geometric vector $\dot{\omega}_{pk}$ varies, but it does so in a perfectly known way. Therefore for each of the $\dot{\mathbf{S}}_p$ components, we have a set of $4K_p$ unknown real scalars. The phase of A_{kj} for a given k, is the same on all sensors and expresses the relative phase of the k^{th} CPW building $\dot{\mathbf{S}}_p$.

Since an $\dot{\mathbf{S}}_p$ cannot be determined better than within a phase factor, we could, in principle take an arbitrary phase for one CPW, for example A_{1p}. This would reduce by one the number of unknown real scalars. However we shall not do so, because we are going to compare this parametric expression of $\dot{\mathbf{S}}_p$ with that given by (4.24) or (4.25). These two expressions must only differ by an unknown phase factor since they each represent the same $\dot{\mathbf{S}}_p$ except for the phase. Thus we must maintain a free phase factor for A_{1p} to represent any possible phase difference between the two expressions for the same $\dot{\mathbf{S}}_p$. As a result the number of free real scalars in the CPW model is confirmed at $4K_p$. The total of CPW for all the $\dot{\mathbf{S}}_p$ is K with

$$K = \sum_{p=1}^{P} K_p, \tag{4.32}$$

and the problem is to assess the value of K by comparison with (4.24). The latter, as we remember, already involves $P(P-1)$ free real scalars.

4.12

At the end of section 4.6, we have already noticed that the n^{th} component of $\dot{\mathbf{S}}_p$ from (4.24) involves all the components of $\dot{\mathbf{Z}}_p$ and therefore, the same $P(P-1)$ free real scalars of matrix $\dot{\mathbf{Z}}$ (obvious in the $P=2$ example of section 4.8). Now the comparison, for the same $\dot{\mathbf{S}}_p$ of the CSDM derived expression and the CPW model yields:

1 vector equation
N complex scalar equations
$2N$ real scalar equations.

For the whole set of $\dot{\mathbf{S}}_p$ we have $2NP$ real scalar equations. The number of free real scalars to be derived from this set of equations is $4K - P(P-1)$

and the value of K is given by $4K + P(P-1) = 2NP$, or solving for K,

$$K = \frac{NP}{2} - \frac{P(P-1)}{4}. \tag{4.33}$$

Therefore for a given P, the number of available CPW is obviously growing with N (the array power) at the rate of P waves for two more sensors, hence, one more wave per wavefront.

Perhaps less obvious, *K is also growing with the number of sources P*, within the limits of its possible variation for a any given N. Indeed the P derivative of the right side of (4.33) is $(N-P)/2 + 1/4$, a positive definite number. It is an important aspect of this assessment that the whole set of source vectors is described better when there are more sources. Later, we shall come back to this remark. Nevertheless the mean value of CPW *per source* decreases as P grows. Though we have mentioned that the CPW may not be (and cannot always be) equally distributed, a rough estimate of the power of the array per source vector is given by the integer nearest to

$$\frac{K}{P} = \frac{N}{2} - \frac{P-1}{4}. \tag{4.34}$$

This number decreases with P, and the P derivative of the right side is $-1/4$ which means that for 4 *more sources* the description of each individual wavefront is reduced *by only one* CPW, an extremely slow rate.

Before giving examples, we shall first make another small assumption; *N is assumed to be an even number*. Considering arrays of 60 sensors or more this is anything but stringent. Equation (4.33) reads

$$K = \frac{N}{2}P - \frac{P(P-1)}{4}, \tag{4.35}$$

with $N/2$ an integer. For K to also be an integer, the even number $P(P-1)$ must be a multiple of 4.

Let us see a few examples where this condition is fulfilled and afterwards what happens when it is not:

1. $N = 60$, $P = 20$ hence $K = 505$
   The whole set of wavefronts is represented by 505 CPW. There cannot be the same number of wavefronts per source, but the mean power of the array can be represented by the integer nearest to $505/20$, namely 25.

2. $N = 62$, $P = 20$ hence $K = 525$
   Adding 2 more sensors (N an even number), the whole set of wavefronts are now described by 20 more CPW, one per source. Indeed the integer nearest to $525/20$ is 26.

3. $N = 60$, $P = 24$ hence $K = 582$
   Coming back to 60 sensors but with 4 more sources as compared with case 1, the global situation is described by 77 more CPW. More sources give more information as could be expected. Nevertheless, per source, the integer nearest to $582/24$ is now 24. It is remarkable that 4 more sources have stolen only one CPW, as compared to case 1, for the descripton of a particular source vector.

4. $N = 60$, $P = 22$ hence $K = 545 = 1/2$
   Let us take now an intermediate example where the even number $P(P - 1)$ is not a multiple of 4. For this case the integer nearest to $545/22$ is either 24 or 25 and the power of an individual wavefront description is about the same as in cases 1 and 3.

What is the meaning of a "half-CPW"? It means that 545 CPW is a bit too much and *actually brings into the system of real scalar equations* too many free real scalars for the number of equations. One of the 545 CPW has become "half-known," depending on only 2 free real scalars instead of 4. *We then need some additional a priori information.* One can build *one* of the 545 CPW with a known direction vector, $\vec{a}$, which is not frequency dependent, a *true* plane wave. The complex amplitude remains undefined. To choose the direction we can rely on any loose a priori idea we might have on the significant presence of a particular plane wave in the description of one of the wavefronts. It is not very critical. An error would be revealed by a small corresponding amplitude in the final computed solution. One could say that the stringent assumption of chapter 2 has now been submerged as a very small part within a flexible solution.

4.13

The number of free real scalars available to describe one wavefront is

$$\frac{4K}{P} - 2N - P + 1. \tag{4.36}$$

It does not seem preposterous to say this number is "large." In the previous examples it is about 100, more precisely it is:

101 in case 1,
105 in case 2,
97 in case 3,
99 in case 4.

This sort of assessment does not help us at all in determining the clustering of these scalars for a sequence of coherent CPW. But it does help us obtain a rough evaluation of the parametric processing power. For an efficient 60 sensor array, it seems sensible to associate a number of "perfect" sources between 10 and 40. The comparable figures are then:

111 free scalars for 10 sources,
81 free scalars for 40 sources,

which exhibit a ratio of only 0.73 scalars for 4 times more sources. Is it going too far to say there is apparently never a shortage of free scalars and their number varies only slowly with the number of sources? Both features are those of a good parametric representation.

Nevertheless it must be emphasized that in order to process a set of 10 sources according to the procedure in section 4.4, we must first find a parametric noise model (no wavefront) able to "eat" 50 eigenvalues out of the 60 x 60 data matrix $\ddot{\mathbf{C}}_r$ and then get the source matrix $\ddot{\mathbf{C}}_s$. Of course a large array might be kind enough to make the subspace division nearly obvious depending on the eigenvalue spread. Still the determination of P is the weak point of the method, even though we might have some clues (doubtful). All of these difficulties arise from trying to avoid faith in stringent wavefront assumptions. But the assumptions we have made *are relatively weak*. Instead of direct statements about the $\vec{\mathbf{S}}_p$ themselves, we make indirect statements about eigenvalues, which are each only a small part of some $\vec{\mathbf{S}}_p$ component [see equation (4.25)]. Also we have not yet used either the frequency dimension or any experimental data about the spreading of the eigenvalues obtained under controlled conditions.

The analysis of the data matrix $\ddot{\mathbf{C}}_r$ must be performed in hundreds of frequency bins. Certainly in an operational situation P is not expected to be the same everywhere within the whole bandwidth. Different sources may transmit components in various different sub-bands. Nevertheless,

when following the same eigenvalue across a few neighboring bins, some stability or smooth variation should be observed, especially for the large bandwidth sources considered here. Fast, erratic fluctuations should disqualify an eigenvalue as being representative of the set of $\dot{\mathbf{S}}_p$.

On the other hand, experiments with P controlled sources are needed (and are already being done) to get some data about the actual spreading of the eigenvalues and the reliability of various methods of estimating the value of P (as if it were unknown). As small as we have left the part of estimating P, it must be backed by experimental data even though there is no universal and definitive criterion to be expected. In short, the behavior of an eigenvalue across a few adjacent frequency bins should contribute to its acceptance as being representative of a set of source vectors and therefore indicative of the number of sources. This is apparently the corner stone of the eigenvalue approach.

The a posteriori verification of a CPW as a true plane wave may be tested by the stability of its direction vector $\dot{a}$ from one frequency bin to another. Probably only a few vectors will qualify for true plane-wave status. But none of this invalidates the validity of the parametric representation. In fact, true plane waves are possibly of interest only for a fast, obvious solution to the localization problem. Again, across the bandwidth, we should observe no erratic variations of the $\dot{a}_{pk}$ and A_{pk}, bound to source vector $\dot{\mathbf{S}}_p$, which are liable to call into question either the $\dot{\mathbf{S}}_p$ or the number of source vectors. Indeed, *everything relies on the estimation of* P.

A huge computational effort, probably not now available in real time, is the cost to be paid for the implementation of this process. The basic mathematical problem to be solved is governed by (4.24) or (4.25), and (4.31), expressing the source vector $\dot{\mathbf{S}}_p$ in two different ways. When solved it yields the source vector $\vec{\mathbf{S}}_p$ along with the apparent spectral density, $d_p = \dot{\mathbf{S}}_p^{\dagger}\dot{\mathbf{S}}_p$, and the wavefront $\dot{\mathbf{F}}_p = d_p^{\frac{1}{2}}\vec{\mathbf{S}}_p$. So in principle the problem of source identification with unknown wavefronts is solved within the sense given to the term "identification" at the beginning of this paper. Let us now remember that since the process is adaptive, the whole solution is valid for a time duration T and liable to be refreshed every T seconds or sooner. As in any rich parametric representation with such enormous flexibility, the geometric and physical significance of so many free parameters is not very important.

4.14

We could stop this presentation here since we are not primarily interested in localization. Nevertheless a few remarks should be made concerning the relationship between what we have done above and the process of localization. First, whatever localization method is to be chosen, it is good to have a fine description of the source vectors. Second, localization is not separable from some description of the propagation medium. In fact the source vectors themselves *are* a description of the medium since *each vector depends on the same medium model as well* as the coordinates and spectral density of the source. We now have the source vectors which we have been trying to describe while using a minimum of assumptions consistent with the array size.

Our present goal is to use the same guide-lines for localization. We seek free parameters which describe the source vectors taking into account the fact that the source vectors are known functions of frequency. We shall approach this problem only in a fairly superficial manner. It must be made clear that what follows is neither a method nor even a theory. Instead it is a conceptual assessment of the free parameters liable to be involved in a parametric description of the medium.

We have already assumed that all sensors exhibit the same flat frequency response and are approximately point-like with a known postion given by a 3-D vector, $\vec{\mathbf{r}}_n$ representing the three coordinates. We now additionally assume there is no calibration problem and therefore, the symbol $\vec{\mathbf{S}}_p$ which has been used to represent a source vector in terms of sensor output elecric voltage, can also be used to represent the sound pressure.

To include the principle of reciprocity, the sound field of a normalized amplitude source, is given by a Green's function

$$\vec{\mathbf{G}}_{pn} = \vec{\mathbf{G}}\,(x_p, y_p, z_p, \vec{\mathbf{r}}_n, a_1, \ldots a_q, \ldots a_Q)\,,$$

which is symetric with respect to the source coordinates x_p, y_p, z_p and the coordinates $\vec{\mathbf{r}}_n$ of the n^{th} sensor; the parameters $a_1, \ldots a_q, \ldots a_Q$ represent a model of the medium. Building the Green's function is generally easy for very simple conditions such as a perfectly reflecting surface or a bottom reflection with a known impedance. The parameters a_q are then very few and known a priori. But the Green's function rapidly becomes very complex for the more realistic fields associated with specific propagation conditions.

e.g. diffraction phenomena etc. So usually the parameters and the form of the Green's function are obtained from an accepted medium model. Most of the time such models are over simplified representations of the real world which only describe the basic outline of an actual acoustic field.

The point of view here is a bit different. We conjecture that it might be possible to build a general Green's function which involves unknown free parameters. We then try to determine how many parameters are needed for compatibility with the array size and the recently acquired knowledge of the source vectors. This is important because this number of available parameters, Q, determines the complexity of the Green's function which can be built. The challenge facing model makers and fast real time computer specialists is the actual building of such flexible Green's functions.

Parameters a_q may be real or complex, frequency dependent or not. But whatever they might be, they are the same for all sources and sensors. Only the source coordinates and the sensor coordinates vary with source and/or sensor position with the latter presumably known. Then the p^{th} source with an unknown complex frequency dependent amplitude Δ_p yields on the n^{th} sensor, an acoustic field $\Delta_p \vec{\mathbf{G}}_{pn}$ which is precisely the n^{th} component of vector $\vec{\mathbf{S}}_p$ now known from (4.25). We must remember that $\vec{\mathbf{S}}_p$ can be known only up to an arbitrary phase factor which is that of Δ_p. Also remember, the *true* spectral density of the source mentioned in chapter 3 is $d_t = |\Delta|^2$ while $d = \vec{\mathbf{S}}^\dagger \vec{\mathbf{S}}$ was only the apparent spectral density as seen through the array. For a given $\vec{\mathbf{S}}_p$ we have N complex scalar relations of the type

$$\Delta_p \vec{\mathbf{G}}_{pn} = s_{pn}. \tag{4.37}$$

This relation involves 4 parameters bound to the source: $\{\Delta_{1p}, x_p, y_p, z_p\}$ which happen to be real scalars, and Q parameters bound to the medium. Considering the whole set of P source vectors, we have:

NP scalar relations,
$4P$ parameters for the sources,
Q parameters for the medium.

The system is balanced when:

$$\begin{aligned} 4P + Q &= NP, \\ Q &= (N-4)P. \end{aligned} \tag{4.38}$$

So, the number of parameters required by the general Green's function of the medium grows with N and P as expected. Nevertheless the full proportionality to P in (4.38) is somewhat misleading. It comes from the fact we are supposed to have first solved for the source vectors, but these vectors are already an implicit description of the medium. Actually, the number Q grows with P at about the same rate as the number of CPW in section 4.12. So we can say that *more sources shed more "light" on the medium description* as could be reasonably expected. All of this *suggests the use of sources for studying the medium.*

From (4.38), the case $N = 60$, $P = 20$ yields $Q = 1120$ which may seem an enormous descriptive capacity. But we must remember that we deal with *scalars* when describing complex propagation phenomena. For the localization problem we may only be interested in the source parameters. So we can "eliminate" the medium description from (4.37) while solving for the source coordinates and spectral density. There is no formal guarantee the coordinates will be perfectly constant from one frequency bin to another. Only the ultimate fitness of the model to the real situation might make it so as a first approximation. But this is also the case for any other passive localization process. The question of whether or not this process is accurate to the limits of its high degree of complexity, is a matter of further experiments.

Opposite to the localization problem, we have the *experimental study of the medium* with sources controlled in both position and spectral density. Individualization of source vectors may result from transmissions in nonoverlapping sub-bands or from the disentangling procedure suggested in sections 4.11–4.13. Once the source vectors are obtained, NP free parameters are available to describe the model. Reiterating a remark already made about the continuous plane waves and large parametric descriptions, the physical meaning of a particular parameter may be less critial when there are more of them. Of course, much further study will be necessary to confirm this possibility.

4.15

As a conclusion for chapter 4, let us say we have at last discarded any a priori knowledge about a source vector except for the *existence* of a wavefront,

tested by the fact that the corresponding CSDM is a rank-one matrix. This is the minimal requirement for spatial processing to be possible. Extensive experimentation at sea is recommended in order to determine the extent to which this last assumption can be trusted across various aperture sizes. A signal is any source with a wavefront, a so-called "perfect" source. Noise is everything else. We are doomed to simultaneously determine all the signal source vectors in a *sort of imaging process.*

In principle the data CSDM measured at the sensor outputs, $\ddot{\mathbf{C}}_r$, involves several "perfect" sources plus noise. The CSDM of P noncorrelated perfect sources is $\ddot{\mathbf{C}}_s$, a rank-P matrix. If P were larger than N there would be no way to disentangle the perfect-source wavefronts from each other (even the knowledge of $\ddot{\mathbf{C}}_s$ is not enough). For a solution to be possible the array has to be large enough for the data CSDM to be the sum of the perfect sources CSDM and of the noise CSDM, $\ddot{\mathbf{C}}_r = \ddot{\mathbf{C}}_s + \ddot{\mathbf{M}}$, where the noise is uncorrelated with any perfect source. From the experimental data of $\ddot{\mathbf{C}}_r$ the first problem is to derive matrix $\ddot{\mathbf{C}}_s$ and therefore the number P of perfect sources.

This could be done accurately at the cost of risky betting about the wavefronts if matrix $\ddot{\mathbf{C}}_s$ is known to have a specific form. Although *this* is sometimes done in the literature, we have discarded such an approach. Instead, we prefer trying to eliminate $\ddot{\mathbf{M}}$. Through a careful examination of the eigenvalues of matrix $\ddot{\mathbf{C}}_r$, we use experience to conjecture how many of the largest eigenvalues represent source vectors, thus determining the value of P. The method is backed by two remarks:

1. Guessing about the eigenvalues is less risky than guessing about the source vectors; since every source vector depends on all the eligible eigenvalues, an error in one of them is only an error within a small part of the final source vector.

2. Guessing is less and less risky when the array grows in size.

A "large," and therefore efficient array has a wide span of eigenvalues, from the largest to the smallest, belonging to the data CSDM with the smallest *decreasing* as N increases. Thus a comparatively vanishing smallest eigenvalue is a sign that the array is sufficiently large to surmount the complexity of the acoustic field which is supposedly being described correctly. Experiments at sea are required for checking the extent to which real arrays fulfill

this condition under operational conditions, or conversely, what size array might be needed.

If the span of the eigenvalues happens to be only a guideline for the choice of P, other clues have to be used, since we must separate matrix $\ddot{\mathbf{C}}_r$ from matrix $\ddot{\mathbf{M}}$. As described here, this step is a slightly generalized version of some methods suggested in the literature. It consists in representing $\ddot{\mathbf{M}}$ with a reliable general model including free parameters which are adjusted to bring matrix $\ddot{\mathbf{C}}_s$ to rank P. If the noise model should prove to be highly reliable (which may well happen for fixed arrays), the choice of P may be affected. When this happens the free parameter adjustment brings matrix $\ddot{\mathbf{C}}_s$ to the minimum possible rank compatible with the noise model.

Besides the existence of the wavefront for the perfect sources, this parametric noise model is another thing which must be restored from total ignorance. Fortunately the final choice for P is not very critical, and is subject to a posteriori verification once the source vectors are determined. Having determined the number of " perfect" sources, P, and matrix $\ddot{\mathbf{C}}_s$, the next problem is to separate the wavefronts from each other. Apparently it is impossible to get the source vectors without additional assumptions, and the information needed corresponds to the degrees of freedom in an arbitrary $(P \cdot P)$ unitary matrix. Any larger constraint would assign matrix $\ddot{\mathbf{C}}_s$ specific features which might conflict with the experimental data from which $\ddot{\mathbf{C}}_s$ was determined. While less constraints would leave the problem unsolved.

Actually we are confronted with an expression for the source vectors involving the genuine matrix $\ddot{\mathbf{C}}_s$ (its eigenvectors and eigenvlaues) plus some free parameters corresponding to the degrees of freedom mentioned above in (1.21) or (1.25). We must then impose constraints which are just fitted to the degrees of freedom left in this "CSDM-derived" source vector expression. The constraint we choose requires each source vector to be expressed as a finite coherent sum of unknown "conditional plane waves" CPW, with the number of plane waves exactly satisfying the available degrees of freedom. So we get a "CPW-derived" expression for the source vectors in such a way that when both equations are satisfied for the whole set of source vectors, we find a sufficient number of scalar equations to compute all the free scalars in both expressions.

The CPW is a flexible parametric expression which occasionally proves to be a plane wave across the useful bandwidth. Although difficult to carry

out, this way of determining the source vectors has several advantages:

1. It preserves the experimental data and depends only on the fitness of a parametric noise model having the flexibility usually associated with adaptivity.

2. It only imposes constraints on the source vectors which are exactly compatible with the array size and number of sources. With the source vectors at hand, an approach to the localization problem is briefly outlined.

Chapter 5

GENERAL CONCLUSION

In chapter 2, while discussing the evolution of increasingly complex spatial processors, we introduced the concept of Directivity and its natural implementation via classical fixed beamformers. This classical approach requires two stringent concepts:

1. A plane wave signal coming from a given direction.

2. An extremely specific and unnatural noise model, based on perfect isotropy and *the absence of correlation between any two nonoverlapping* elementary solid angles.

The above leaves an insufficient amount of information to exploit for the "identification" of a source by the determination of its wavefront and spectral density. Here the wavefront is already known and the best signal to noise ratio is obtained for the particular sensor combination which maximizes the directivity factor.

When trying to go further, through an estimation of the "noise alone" on the output of some other appropriate sensor combination, we fall on the rigid concept which grants no specific status to remote sources transmitting plane waves from directions other than the signal direction; remote sources also produce noises, or "unwanted signals" on the "beam" which we would like to clear of anything but *the* signal. Nevertheless under some assumptions discussed in chapter 3, it is possible to reach the notion of "noise alone" and its estimation. But all such procedures rely on an excess of confidence in the highly artificial notion of omnidirectional noise. In fact

the Directivity concept does not provide reliable removal of the noise spectral density from the raw data, and is thus unable to reach the realm of the so called "hyper-resolution" systems where the signal spectral density *is supposed to be totally rid of noise.*

The next approach to spatial processing is the adaptive beamformer for which the signal wavefront is still known (its spectral density needs to be identified) but for which *no* other assumption is made about the unwanted noise. The same status also covers sources with a wavefront (but different from the signal's) and diffuse sources with no wavefront; it is all unwanted noise. Given the signal wavefront, building a sensor combination which preserves this signal (a "signal all pass filter") is easy, but the output still contains some unwanted noise. Knowledge of the signal wavefront also makes it possible to build sensor combinations which do not let the signal through, even when present. These combinations are the "Signal Free References" (SFR). Generally, $(N \quad 1)$ SFRs are available from an N-sensor array, and there are several ways to render them uncorrelated using processes which are the physical analog of the diagonalization of the SFR cross-spectral density matrix (CSDM). Nevertheless the SFRs remain correlated with the noise gathered on the signal output.

In this paper a way is described to progressively subtract contributions from the noises that were adaptively derived from the SFRs. Strange but true, these subtractions are both those of noises and their spectral densities. So that, in a repetitive operation exhausting all SFRs the residual noise left in the (unaltered) signal company, has been brought to a minimum. This residual noise cannot be estimated directly since we know neither whether nor when a signal may be present. Nevertheless and this is a success of the adaptive beamforming theory—it is possible to show that *only one scalar* relation between the elements of the CSDM is needed to obtain an estimate of the minimized noise above, thus reaching the realm of hyper-resolution systems. Such a relation must be derived from intuition or more reliably bound to a particular situation. For a large matrix this is a "weak" assumption which does not appear to be too critical.

Such an assumption can also be considered as automatically pertaining to "large arrays." In chapters 3 and 4 a large array is a one whose size (or complexity, or number N) is large enough *to surmount the complexity* of the number of sources plus the complexity of the propagation medium. The test of this capacity is derived from the examination of the span of

the CSDM eigenvalues, in every frequency bin. It must be noticed that more sensors lowers the value of the smallest eigenvalue and is liable to widen the span of values. It is important for the smallest eigenvalues to be much smaller than the largest. We then can consider just neglecting this smallest eigenvalue, which immediately provides the wanted scalar relation mentioned above and yields the signal, in principle, rid of noise inasmuch as second order statistics are concerned. This "large array condition" remains to be experimentally checked on actual arrays in operational conditions. Such a check must done before stating that the array correctly represents the medium and sources, and therefore is an efficient array.

Noise subtraction techniques and noise cancelling techniques are all particular branches of this presentation as is the removal of self-noise from hull-mounted arrays. Although one may be enthusiastic about the idea of nearly achieving hyper-resolution while dropping all assumptions concerning unwanted noise, some caution is in order because the process is very sensitive to the shape of the signal wavefront. A small deviation of a signal away from its expected wavefront unpredictably alters the output signal spectral density, a quantity we had been careful to preserve. Making the process more robust is possible but this incurs a loss in the noise reduction capacity. Besides it is impossible to be prepared for all the possible wavefronts describable in an N-dimensional space (a generalization of beamforming). To do so would be both fantastically complex and a true waste, since we are only considering "large arrays," with far fewer sources than sensors.

Finding a way to process signals with unknown wavefronts forces us to restore some "weak" assumption--the weakest possible. We base our assumption on the observation that a source with an unknown wavefront *must nevertheless have some "existing" wavefront.* Whether this condition is fulfilled in the real world will be decided by experimental studies which determine whether an "interesting" source of the type we might have to identify exhibits a rank-one CSDM regardless of the wavefront. This is, along with the eigenvalue span, the corner-stone of spatial processing and its ultimate justification. One may expect the corner-stone is firmly attached to reality, especially for large fixed arrays.

On the other hand, the global CSDM of several "perfect" sources, each possessing a wavefront, is not by itself sufficient for the separation of several source vectors. One can only say *the rank of this "perfect sources CSDM" is the number P of sources.* Furthermore noise sources with a diffuse structure

(not granted a wavefront) add their own full rank CSDM to yield the only accessible matrix, the raw data CSDM. The procedure to be considered is then a two-step one.

The first step consists in trying to eliminate, as much as possible, the noise CSDM from the raw data CSDM. To do so we need a reasonably general model of the noise CSDM, which is made flexible enough through the use of free parameters (in principle $(N - P)$ parameters if P is already chosen), to be fitted to the widest category of situations. Then the free parameters are varied to reduce the rank of the difference between the raw data CSDM and the noise CSDM from N to P. The remaining difference matrix is expected to be the source CSDM.

While the number of P sources might usually be derived from an examination of the set of eigenvalues in the raw data CSDM, under some conditions the noise model may be reliable enough (especially for fixed arrays) to either impose, or at least influence the choice of P. In fact the choice of P cannot be strictly determined precisely because we have made no assumption about the wavefronts. On the other hand, this is not a critical choice since it is subject to a posteriori verification.

In the second step, P is now known and the second order statistics of the noise have been eliminated. What remains is to disentangle the source vectors from their $(N \cdot P)$ CSDM. Unfortunately this disentanglement cannot be done without some additional assumptions because there are an unlimited number of solutions for source vectors directly derived from the CSDM. These solutions depend on a set of undetermined parameters corresponding to the degree of freedom of a $(P \cdot P)$ unitary matrix. As a consequence we have to impose some "weak" constraint on the source vectors, so that the free parameters become computable thus allowing a solution to emerge.

The constraint suggested here is a parametric representation of a source vector as a coherent sum of a *limited* number of "conditional plane waves," CPW, which are not exactly plane, but whose geometry is only subject to a posteriori control while the parametric value remains constant. The limited number of CPW per source vector and their flexibility in terms of physics, seem to justify the claim for the assumption being "weak."

The constraints on the CPW in this source vector model are chosen in such a way that the number of free scalars on both sides exactly match the number of relations obtained when the "CSDM derived" expression is

confronted with the same source vector. Most of the assumptions about the source vectors which have been proposed in the literature are far from this point of balance, and actually impose a specific form onto an already known experimental CSDM.

Our notion of CPW seems to be acceptable as a weak assumption because in a rich parametric representation, the physical definition of a given parameter tends to be blurry without loss of the parametric value. Also, this might be true for a parametric localization process for the assessment of available free parameters in the description of the medium, such as the one briefly sketched at the end of chapter 4.

Appendix A

Classical Relation between Cross-Spectral Densities Before and After Filtering

The following well known rule is one of the basic tools of the mathematical presentation:

$$B_1^i \longrightarrow \boxed{\Phi_1} \longrightarrow B_1^0$$

$$B_2^i \longrightarrow \boxed{\Phi_2} \longrightarrow B_2^0$$

Given two input noises B_1^i and B_2^i respectively filtered with frequency responses Φ_1 and Φ_2 which deliver two output noises B_1^0 and B_2^0, we have

$$\mathrm{csd}\left[B_1^0 B_2^0\right] = \mathrm{csd}\left[B_1^i B_2^i\right] \Phi_1 \Phi_2^*, \tag{A.1}$$

where csd denotes cross spectral density and sd spectral density. In the particular case where B_1^i and B_2^i are the same noise then,

$$\mathrm{csd}\left[B_1^0 B_2^0\right] = \mathrm{sd}\left[B_i\right] \Phi_1 \Phi_2^*. \tag{A.2}$$

Appendix B

Matrix of Filters and Related Equations

As a rehersal with the CSDM and a matrix of filters let us consider N original noises:

$$B_1, B_2, \ldots B_n, \ldots B_m, \ldots B_N;$$

with a CSDM $\bar{\bar{\mathbf{C}}}$ whose elements are c_{nm} (n^{th} row, m^{th} column). Let these elements be processed through a *matrix of filters* as shown in figure B.1. The frequency response of a *filter is* g_{pn} *where:*

1. The first subscript refers to the p^{th} output noise B_p^0 which is filtered by the p^{th} row of filters.

2. The second subscript refers to the n^{th} input noise B_n which feeds all the filters of the n^{th} column of filters.

So the matrix of filters $\bar{\bar{\mathbf{G}}}$ appears as a matrix is usually written—with elements g_{pn} in the p^{th} row and the n^{th} column. Using the symbol $\otimes$ for "filtered by", we have:

$$B_p^0 \qquad (B_1 \otimes g_{p1}) + \ldots + (B_n \otimes g_{pn}) + \ldots + (B_m \otimes g_{pm}) + \ldots + (B_N \otimes g_{pN}),$$
$$B_q^0 \qquad (B_1 \otimes g_{q1}) + \ldots + (B_n \otimes g_{qn}) + \ldots + (B_m \otimes g_{qm}) + \ldots + (B_N \otimes g_{qN}).$$

Using the rule from appendix A, the cross-spectral density (csd) between the two noises is,

$$\text{csd}\left[B_p^0 B_q^0\right] \qquad g_{pq}$$

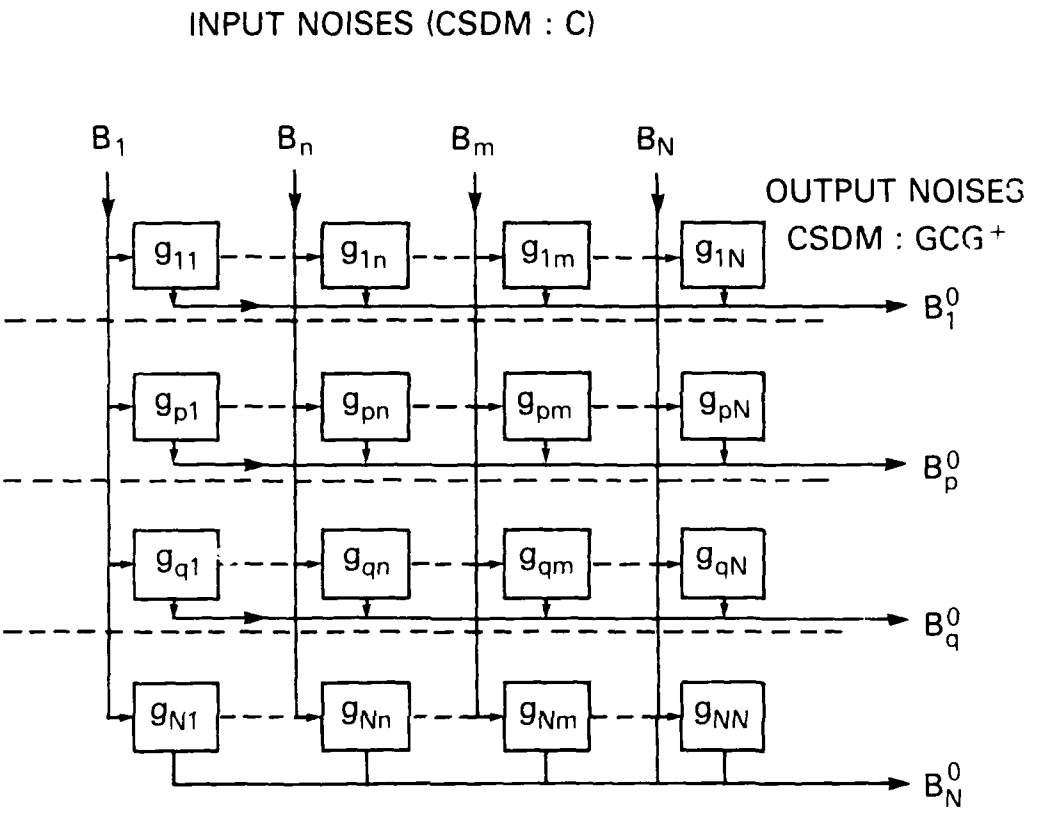


Figure B.1: Matrix of Linear Filters.

$$y_{pq} = \sum_{n=1}^{N} \sum_{m=1}^{N} g_{pn} c_{nm} g^{*}_{qm}, \tag{B.1}$$

which is an element of the output CSDM $\ddot{\mathbf{Y}}$.

Now g^{*}_{qm} is the element γ_{mq} of the matrix $\ddot{\mathbf{G}}^{\dagger}$, so that y_{pq} can be written

$$y_{pq} = \sum_{n=1}^{N} \sum_{m=1}^{N} g_{pn} c_{nm} \gamma_{mq} \tag{B.2}$$

which is exactly the mathematical expression of the element for a product of three matrices. Indeed an element of $\ddot{\mathbf{H}} = \ddot{\mathbf{G}}\ddot{\mathbf{C}}\ddot{\mathbf{D}}$ is

$$h_{pq} = \sum_{n=1}^{N} \sum_{m=1}^{N} g_{pn} c_{nm} d_{mq}, \tag{B.3}$$

Comparing (B3) with (B2) clearly shows

$$\ddot{\mathbf{Y}} = \bar{\bar{\mathbf{G}}}\ddot{\mathbf{C}}\bar{\bar{\mathbf{G}}}^{\dagger}. \tag{B.4}$$

The spectral density of B_p^0 is

$$\begin{aligned} \mathrm{sd}\left[B_p^0\right] &= y_{pp} \\ &= \sum_{n=1}^{N}\sum_{m=1}^{N} g_{pn} c_{nm} g_{pm}^{*}. \end{aligned} \tag{B.5}$$

Considering the sequence of g_{pn} with fixed p as the row vector $\bar{\mathbf{g}}_p$, the sequence of g_{pm}^{*} becomes the column vector $\bar{\mathbf{g}}_p^{\dagger}$. Therefore

$$y_{pp} = \bar{\mathbf{g}}_p \ddot{\mathbf{C}} \bar{\mathbf{g}}_p^{\dagger}. \tag{B.6}$$

the quadratic form of vector $\bar{\mathbf{g}}_p$ (row matrix) through matrix $\ddot{\mathbf{C}}$.

To follow the convention made in section 3.3, we should represent the set of filters $\bar{\mathbf{g}}_p$ by a column vector $\bar{\mathbf{V}} = \bar{\mathbf{g}}_p^{\dagger}$ so that $y_{pp} = \bar{\mathbf{V}}^{\dagger}\ddot{\mathbf{C}}\bar{\mathbf{V}}$ as in section 3.4 equation (3.13), and section 3.5.

Appendix C

Classical Factorization of a CSDM - Normalized Eigenvectors

A positive definite, Hermitian matrix like the CSDM $\ddot{\mathbf{C}}$ is known to factor into the product

$$\ddot{\mathbf{C}} = \ddot{\mathbf{U}}\ddot{\mathbf{E}}\ddot{\mathbf{U}}^{\dagger}, \tag{C.1}$$

where:

$\ddot{\mathbf{E}}$ is the diagonal matrix of the eigenvlues, $\epsilon_1, \epsilon_2, \ldots \epsilon_N$.

$\ddot{\mathbf{U}}$ is the unitary matrix whose *columns* are the *normalized* eigenvectors of $\ddot{\mathbf{C}}$: $\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2, \ldots \vec{\mathbf{u}}_N$; placed in the same order as the eigenvalues in $\ddot{\mathbf{E}}$.

Because these eigenvectors are normalized and orthogonal to each other:

$$\begin{aligned} \vec{\mathbf{u}}_p^{\dagger}\vec{\mathbf{u}}_p &= 1, \\ \vec{\mathbf{u}}_p^{\dagger}\vec{\mathbf{u}}_q &= 0, \quad (p \neq q) \\ \ddot{\mathbf{U}}^{\dagger}\ddot{\mathbf{U}} &= \ddot{\mathbf{I}}_N, \quad \text{the identity matrix.} \end{aligned}$$

One may rapidly check that equation (C.1) is the right one when the *columns* of $\ddot{\mathbf{U}}$ are the eigenvectors. The definition of $\vec{\mathbf{u}}_p$ is,

$$\ddot{\mathbf{C}}\vec{\mathbf{u}}_p = \epsilon_p\vec{\mathbf{u}}_p.$$

If equation (C.1) is true we must have

$$\bar{\bar{\mathbf{U}}}\bar{\bar{\mathbf{E}}}\bar{\bar{\mathbf{U}}}^{\dagger}\vec{\mathbf{u}}_p = \epsilon_p\vec{\mathbf{u}}_p. \tag{C.2}$$

We check this by multiplying out the left hand side. Beginning with $\bar{\bar{\mathbf{U}}}^{\dagger}\vec{\mathbf{u}}_p$, and noticing the rows of $\bar{\bar{\mathbf{U}}}^{\dagger}$ are $\vec{\mathbf{u}}_p^{\dagger}$, we end up with a column vector $\vec{\mathbf{v}}$ where all the elements are null except the p^{th} whose value is one. Pursuing with $\bar{\bar{\mathbf{E}}}\vec{\mathbf{v}}$, we get a column vector with all elements null except the p^{th} whose value is ϵ_p. Finally in the matrix product

$$\bar{\bar{\mathbf{U}}}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \epsilon_p \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

only the elements of the p^{th} column of $\bar{\bar{\mathbf{U}}}$ are scaled by the common factor ϵ_p. The result is then $\epsilon_p\vec{\mathbf{u}}_p$ as expected from (C.2).

Appendix D

Orthogonal Images of a Set of Correlated Noises

We now know that through the matrix of filters $\ddot{\mathbf{G}}$, the former CSDM $\ddot{\mathbf{C}}$ has become $\ddot{\mathbf{G}}\ddot{\mathbf{C}}\ddot{\mathbf{G}}^{\dagger}$ which may also be written as $\ddot{\mathbf{G}}\ddot{\mathbf{U}}\ddot{\mathbf{E}}\ddot{\mathbf{U}}^{\dagger}\ddot{\mathbf{G}}^{\dagger}$. If we choose to use a matrix of filters $\ddot{\mathbf{G}} = \ddot{\mathbf{U}}^{\dagger}$, clearly the output matrix becomes

$$\left[\ddot{\mathbf{U}}^{\dagger}\ddot{\mathbf{U}}\right]\ddot{\mathbf{E}}\left[\ddot{\mathbf{U}}^{\dagger}\ddot{\mathbf{U}}\right] = \ddot{\mathbf{E}}, \tag{D.1}$$

because $\ddot{\mathbf{U}}^{\dagger}\ddot{\mathbf{U}} = \ddot{\mathbf{I}}_N$ is the identity matrix. So we have got a set of uncorrelated noises with spectral densities $\epsilon_1, \epsilon_2, \ldots \epsilon_p, \ldots \epsilon_q \ldots \epsilon_N$. This is the physical interpretation of the diagonalization of matrix $\ddot{\mathbf{C}}$. The new independent noises are the "orthogonal images" of the input noises. The choice of $\ddot{\mathbf{G}} = \ddot{\mathbf{U}}^{\dagger}$ shows that a row of filters in $\ddot{\mathbf{G}}$, is the corresponding row of the $\ddot{\mathbf{U}}^{\dagger}$ elements, and therefore the *complex conjugates* of the elements of the corresponding *column* in $\ddot{\mathbf{U}}$. In other words the normalized eigenvector $\ddot{\mathbf{U}}_i$ is the column vector which is representative of the set of filters made of the complex conjugates of its elements. We are therefore faithful to the convention stated in section 3.3.

Appendix E

Canonical Form of a CSDM

There is another canonical form of matrix $\bar{\bar{C}}$ which is of use in chapter 4,

$$\bar{\bar{C}} = \sum_{n=1}^{N} \epsilon_n \left[\vec{u}_n \vec{u}_n^{\dagger} \right]. \tag{E.1}$$

In brackets is the dyadic product of a normalized eigenvector which forms an $(N \times N)$ matrix. It is easy to check the previous equation, for example with the eigenvector $\vec{u}_1$,

$$\bar{\bar{C}}\vec{u}_1 = \sum_{n=1}^{N} \epsilon_n \vec{u}_n \left(\vec{u}_n^{\dagger} \vec{u}_1 \right),$$

where the scalars are between parenthesis. Since

$$\begin{aligned} \vec{u}_1^{\dagger} \vec{u}_1 &= 1, \quad \text{and} \\ \vec{u}_n^{\dagger} \vec{u}_1 &= 0, \quad \text{for } n = 1, \end{aligned}$$

we end up with

$$\bar{\bar{C}}\vec{u}_1 = \epsilon_1 \vec{u}_1,$$

which is the very definition of $\vec{u}_1$. The same then holds true for all $\vec{u}_p$.

Appendix F

Preliminary "Theorem" for Section 4.6

Let us consider P vectors: $\dot{\mathbf{V}}_1, \ldots \dot{\mathbf{V}}_P$, with $P < N$ describing a true P-dimensional subspace (no linear dependence between any pair of vectors) of an N-dimensional vector space. Also consider a diagonal $(N \times N)$ matrix $\ddot{\Pi}$ [see equation (4.111)], where the first P diagonal elements are 1, and the remaining $(N - P)$ others are 0. Now recall equation (4.12) from section 4.6,

$$\ddot{\Pi} = \sum_{p=1}^{P} \dot{\mathbf{V}}_p \dot{\mathbf{V}}_p^{\dagger}.$$

Also recall the vectors $\dot{\mathbf{V}}_p$ have two specific features 1 and 2 given in section 4.6. With the components of $\dot{\mathbf{V}}_p$ given by: $v_{p1}, v_{p2}, \ldots v_{pN}$, the diagonal elements of the matrix $\dot{\mathbf{V}}_p \dot{\mathbf{V}}_p^{\dagger}$ are: $|v_{p1}|^2, \ldots |v_{pN}|^2$. The n^{th} diagonal element of equation (4.12) is

$$\sum_{p=1}^{P} |v_{pn}|^2 .$$

Comparing with matrix $\ddot{\Pi}$ it is clear that if $n > P$ the above quantity is zero. Therefore feature 1: any component of any vector $\dot{\mathbf{V}}_p$ is zero when $n > P$. Only the first P components may be nonzero. The set of vectors $\dot{\mathbf{V}}_p$ describes a subspace spanned by the first P axes in the N dimensional space. As a consequence: $\ddot{\Pi}\dot{\mathbf{V}}_p = \dot{\mathbf{V}}_p$. This may be seen by postmultiplying

equation (4.12) by $\vec{\mathbf{V}}_1$,

$$\vec{\Pi}\vec{\mathbf{V}}_1 = \sum_{p=1}^{P} \vec{\mathbf{V}}_p \left(\vec{\mathbf{V}}_p^{\dagger} \vec{\mathbf{V}}_1 \right).$$

In this relation $\vec{\mathbf{V}}_1$ appears to be a linear combination of the other vectors $\vec{\mathbf{V}}_p$. But this is impossible since they have been taken as describing a true P-dimensional subspace. The only way to comply with this equation is to accept:

$$\vec{\mathbf{V}}_p^{\dagger} \vec{\mathbf{V}}_1 = 0 \quad \text{for } p \neq 1,$$
$$\vec{\mathbf{V}}_1^{\dagger} \vec{\mathbf{V}}_1 = 1.$$

Vector $\vec{\mathbf{V}}_1$ is a normalized vector orthogonal to the others. What is true for $\vec{\mathbf{V}}_1$, is true for any $\vec{\mathbf{V}}_p$. Therefore feature 2: the $\vec{\mathbf{V}}_p$ are normalized vectors orthogonal to each other.

Bibliography

[1] C. Turcat-Latombe, "Détection et Caractérisation de Signaux à Plusieurs Composantes à Partir de la Matrice Interspectrale," Ph.D. Thesis, Institut National Polytechnique Université de Grenoble, Grenoble, France, 19 novembre 1982.

[2] G. Bienvenu and L. Kopp, "Adaptive High Resolution Passive Methods," Proceedings of the First European Signal Processing Conference, EUSIPCO 80, September 16-19, 1980, Lausanne, Switzerland, North Holland Pub., pp. 715-721.

[3] C. Hanna, "Analyse Spectrale et Interspectrale de Signaux par Filtrage Autoregressif," Thesis, Centre d'Etude des Phénomènes Aléatoires, Institut National Polytechnique de Grenoble, Grenoble, France, 1979.

[4] J.L. Lacoume, B. Bouthemy, F. Glangeaud, C. Latombe, and A. Silvent, "Caractérisation par Analyse Interspectrale du Champ d'Ondes Reçues sur un Réseau de Capteurs," Septième Colloque sur le Traitement du Signal et ses Applications, Nice, France, 1979, pp. 80/1-80/7.

[5] S.L. Marple, "A New Autoregressive Spectrum Analysis Algorithm," IEEE Trans. ASSP, **ASSP-28**, 441-454, (1980), USA.

[6] T. Thorvaldsen, "Maximum Entropy Spectral Analysis in Antenna Spatial Filtering," IEEE Trans. Antennas and Propagation, **AP-28(4)**, (1980), USA.

[7] C.L. Byrne and R. Fitzgerald, "A Unifying Model for Spectral Estimation," Proc. Rome Air Development Center Spectrum Estimation Workshop, Rome, NY, USA, October 1979, pp. 157-162.

[8] C.L. Byrne and R. Fitzgerald, "Spectral Estimators that Extend the Maximum Entropy and Maximum Likelihood Methods," SIAM J. Appl. Math., **44(2)**, 425-442, (1984), USA.

[9] C.L. Byrne and R. Fitzgerald, "High-Resolution Beamforming with Oversampled Arrays," J. Acoust. Soc. Am., **74(4)**, 1224-1227, (1983), USA.

[10] R. Fitzgerald and C.L. Byrne, "Extrapolation of Band Limited Signals: A Tutorial," Proceedings of the First European Signal Processing Conference, EUSIPCO 80, September 16-19, 1980, Lausanne, Switzerland, North Holland Pub., pp. 175-180.

[11] A. Papoulis, "A New Algorighm in Spectral Analysis and Band-Limited Extrapolation," IEEE Trans. **CAS-22**, 735-742, (1975), USA.

[12] J.E. Shore, "Minimum Cross-Entropy Spectral Analysis," NRL Memorandum 3921, January 1979, Naval Research Laboratory, Washington, DC, USA.

[13] *J.L. Lacoume, F. Glangeaud, J.Y. Jourdain, and N. Martin*, "Spectral and Interspectral Analysis of Low Frequency Submarine Acoustic Signals Received on an Array of Sensors," Lecture at the Nato Advanced Study Institute (NATO ASI) on Adaptive methods in Underwater Acoustics, 30 July - 10 Aug. 1984, Luneburg, Germany.

[14] V.C. Anderson, "The Deltic Correlator," Technical Memorandum 37, Office of Naval Research, Contract N5 OR 1 - 76 Project Order X, USA.

[15] H. Mermoz, "Modularité du Traitement Adaptatif d'Antenne," Annales des Télécommunications, **29(1-2)**, 43-54, (1974), France.

[16] D. Baudois, W. Kaufman, A. Silvent, "Etude Théorique et Expérimentale du Correlo-Filtre," Huitième Colloque sur le Traitement du Signal et ses Applicaitons, Nice, France, 1-5 juin 1981, pp. 269-274.

[17] G. Bouthemy, "Etude d'un Correlo-Filtre Adaptatif," These de Docteur-Ingénieur, Institut National Polytechnique de Grenoble, Grenoble, France, 8 novembre 1982.

[18] H. Mermoz, "Imagerie, Corrélation et Modèles," Annales des Télécommunications, **31(12)**, 17-36, janvier-février (1976), France.

[19] G. Bienvenu, D. Pillon, "Méthodes Haute Résolution en Ecoute Passive," Huitième Colloque sur le Traitement du Signal et ses Applications, Nice, France, 1-5 juin 1981, pp. 341-347.

[20] G. Bienvenu, "Propriétés Haute Résolution de la Matrice de Corrélation Spatiale," Neuvième Colloque sur le Traitement du Signal et ses Applications, Nice, France, 16-20 mai 1983, pp. 239-245.

[21] N.L. Owsley, J.F. Law, "Dominant Mode Power Spectrum Estimation," Proceedings ICASSP 82, Paris, France, May 3-5 1982, pp. 775-778.

[22] J.Y. Jourdain, "Modèles et Milieux," Neuvième Colloque sur le Traitement du Signal et ses Applications, Nice, France, 16-20 mai 1983, pp. 905-910.

[23] A. Plaisant and G. Bienvenu, "Modèles de Bruit Ambiant en Très Basse Fréquence," Neuvième Colloque sur le Traitement du Signal et ses applications, Nice, France, 16-20 mai 1983, pp 911-917.

[24] H. Mermoz, "Spatial Processing Beyond Adaptive Beamforming," JASA, **70(1)**, 71-79, July (1981), USA.

[25] G. Vezzosi, "Séparation des Fronts d'Ondes par Redressement du Facteur de Choleski de Leur Matrice Interspectrale," Annales des Télécommunications, **37(9-10)**, 425-439, (1982), France.

[26] C. Latombe, F. Glangeaud, and C. Lathuillere, "Traitement Matriciel de Signaux à N Composantes et Etude par Couples," Neuvième Colloque sur le Traitement du Signal et ses Applications, Nice, France, 16-20 mai 1983, pp. 253-258.

[27] L. Kopp and G. Bienvenu, "Détection par les Valeurs Propres de la Matrice Interspectrale," Neuvième Colloque sur le Traitement du Signal et ses Applications, Nice, France, 16-20 mai 1983, pp.265-269.

END

DATED

FILM

8-88

DTIC