

Algebraic Aspects of the Exact Signal Demodulation Problem

E. Hubert* Y. Bouzidi** R. Dagher*** A. Barrau****
A. Quadrat†

* *Safran Tech, France (e-mail: elisa.hubert@safrangroup.com)*

** *Inria Lille, France (e-mail: yacine.bouzidi@inria.fr)*

*** *Inria Chile, Chili (e-mail: roudy.dagher@inria.fr)*

**** *Safran Tech, France (e-mail: axel.barrau@safrangroup.com)*

† *Inria Paris, Ouragan project, IMJ – PRG, Sorbonne University, France (e-mail: alban.quadrat@inria.fr)*

Abstract In this paper we introduce a general class of problems originating from gearbox vibration analysis. Based on a previous work where demodulation was formulated as a matrix approximation problem, we study the specific case applicable to amplitude and phase demodulation. This problem can be rewritten as a polynomial system. Based on algebraic methods such as linear algebra and homological algebra, we shall focus on the characterization of the problem and solve it in the noise-free case.

Keywords: Demodulation, polynomial systems, linear systems, homological algebra

1. INTRODUCTION

Modulation and *demodulation* are basic tools of the theory of signal processing which can be described as varying a parameter of a sine wave, which can be its amplitude, phase or frequency, to encode a message (modulation) or to recover the message from such a modified sine wave (demodulation). Motivated by telecommunications, analogical methods were first developed to perform these operations in the time domain although the theory supporting them was based on spectral considerations.

Although these early solutions have encountered a wide success and are still the standard approach to demodulation, improving them is an active field of research due to both the new possibilities brought by digital signal processing and the specificities of new applications such as monitoring of mechanical systems (McFadden (1986)). The latter work is the initial motivation for the present work. The models encountered in these fields differ from the classical sine wave modulation/demodulation (also called “mono-carrier modulation”) because the energy of the modulated signal is spread over several harmonics (multi-carrier modulation) and filtering out all of them but one would mean losing most of the available information. Note this situation also exists in telecommunications although it is not standard (Friedlander et al. (1995)).

Regarding the new possibilities offered by digital signal processing, the first one is to go from the time to the spectral domain by means of the Fast Fourier Transform (FFT). In the time domain, multi-carrier demodulation amounts to estimating the envelope of a signal. The energy separation algorithm achieving this task uses nonlinear differential operators (Potamianos et al. (1994)) such as the Teager-Kaiser energy operator derived in Kaiser (1990), which also extracts the instantaneous frequency

of modulated signals. The energy separation algorithm has also been improved by the iterative generalized demodulation method proposed in Feng et al. (2011). In the spectral domain, the Hilbert transform provides an estimation of the envelope of a mono-carrier modulated signal, along with its phase if needed (Potamianos et al. (1994)). For multi-carrier demodulation, most approaches are two-steps: the signal is first decomposed into mono-component signals which are then demodulated with one of the previous methods or any variation of them. In Friedlander et al. (1995), the separation is based on band-pass filtering. In Santhanam et al. (2000), the authors take advantage of the periodicity of the signal. The separation is then performed with a periodicity-based algebraic algorithm and the demodulation with an energy-based algorithm (PASED). Another method, called Hilbert-Huang transform, introduced in Huang et al. (1998), analyzes the signal by taking it apart into intrinsic mode functions (IMF) using the empirical mode decomposition. A Hilbert transform is then performed on the obtained IMFs. For non-stationary analysis, wavelet based methods have the advantage of allowing demodulation of signals whose time-frequency distributions are curved paths (Yu et al. (2016)). In mechanical signal processing, multicomponent demodulation algorithms have been adapted to fault diagnosis (Feldman (2011)).

In the work Hubert et al. (2018), the amplitude demodulation was studied as an optimization problem where the carrier/modulation couple fitting the data at best is sought. We showed an unexpected correspondence between this formulation and low-rank matrix approximation, and leveraged it to design a new “optimal demodulation” algorithm based on eigenvector computation. In this paper, we keep following the same route and extend the approach to combine amplitude and phase demodulation. Under

the assumption that the phase variations stay small with respect to 2π , the first contribution of the paper is to show that a matrix reformulation, similar to the one developed in Hubert et al. (2018) can be obtained. It yields a constrained rank 2 approximation problem which, to the authors knowledge, has not been given much attention in the linear algebra literature. Our second contribution is to study and effectively solve this problem by means of algebraic methods.

In Section 2, we first derive a novel formulation of the amplitude and phase demodulation as a constrained low-rank matrix approximation problem. In Section 3, we review standard results on linear algebra that will be of constant use. Finally, Section 4 studies the rank 2 decomposition problem corresponding to the exact phase and amplitude demodulation, and gives explicit necessary and sufficient conditions for the existence of a solution.

2. PROBLEM FORMULATION

2.1 Modulation-based mechanical models

Let $T \in \mathbb{R}_{>0}$ be a non-negative real number and $D(T)$ the set of T -periodic functions of $L^1([0, T])$. According to Fourier analysis, for $c \in D(T)$, we have:

$$c(t) = \sum_{k \in \mathbb{Z}} c_k e^{i2\pi \frac{k}{T} t}, \quad c_k = \frac{1}{T} \int_{-T/2}^{T/2} c(t) e^{-i2\pi \frac{k}{T} t} dt. \quad (1)$$

For $n, N \in \mathbb{N}$, the n^{th} harmonic of c is the term defined by $c_{-n} e^{-i2\pi \frac{n}{T} t} + c_n e^{i2\pi \frac{n}{T} t}$ and we call the *first N harmonics* of c the following truncation of (1):

$$H_n(c)(t) = \sum_{0 \leq |k| \leq N} c_k e^{i2\pi \frac{k}{T} t}.$$

Let $D(T, N)$ be the set formed by all the functions of $D(T)$ whose at most first N harmonics are non-zero, i.e.:

$$D(T, N) = \{f \in D(T) \mid |k| > N : c_k = 0\}.$$

In this paper, we consider a class of time signals arising in the study of *gearbox vibrations* that are of the form

$$s(t) = \sum_{j=1}^q m_j(t) c_j(t), \quad \mathcal{L}(m, c) = 0, \quad (2)$$

where s denotes the output of the sensor, $m = (m_j)_{j=1, \dots, q}$ and $c = (c_j)_{j=1, \dots, q}$ are the unknown components to be estimated which are subjected to certain linear constraints $\mathcal{L}(m, c) = 0$. Let us give a few example of such situations.

Amplitude modulation without overlap is the prototypical example of (2) which can be obtained by considering

$$s(t) = c(t) m(t), \quad (3)$$

with the constraints $m \in D(N_m, T_m)$, $c \in D(N_c, T_c)$, $T = T_m N_m$ for a certain $N_m \in \mathbb{N}$ (resp., $T_c N_c = T$ for a certain $N_c \in \mathbb{N}$), and:

$$\mathcal{L}(m, c) = \begin{cases} c_k = 0, & k \notin \{1, \dots, N_c\}, \\ m_k = 0, & \text{for } |k| \geq f_c/2. \end{cases}$$

Hence, the value of q is 1 and the linear constraints amount to say that c has frequency $f_c = T_c^{-1}$ and the spectral support of m is in $] -f_c/2, f_c/2[$.

Let us now consider a modified version of (3) which encodes the *phase and amplitude modulation* signal:

$$\begin{aligned} s(t) &= c(t + \phi(t)) m(t), \\ m &\in D(N_m, T_m), \quad c \in D(N_c, T_c), \quad f_c = T_c^{-1}, \\ \mathcal{L}(m, c) &= \begin{cases} c_k = 0, & k \notin \{1, \dots, N_c\}, \\ m_k = 0, & \text{for } |k| \geq f_c/2, \\ \phi_k = 0, & \text{for } |k| \geq f_c/2, \end{cases} \end{aligned} \quad (4)$$

A study of this general system is difficult. But under the additional hypothesis that the phase modulation is small

$$\phi(t) \ll 2\pi,$$

we can recast it into the general framework of (2). Indeed, we can perform a first-order expansion w.r.t. ϕ and write:

$$\begin{aligned} c(t + \phi(t)) m(t) &= (c(t) + c'(t) \phi(t)) m(t) \\ &= c(t) m(t) + c'(t) \phi(t) m(t). \end{aligned}$$

Setting $c_1 = c$, $c_2 = c'$, $m_1 = m$, $m_2 = m \phi$, (4) yields:

$$\begin{aligned} s(t) &= c_1(t) m_1(t) + c_2(t) m_2(t), \\ m_1 &\in D(N_m, T_m), \quad m_2 \in D(N_m, T_m), \\ c_1 &\in D(N_c, T_c), \quad c_2 \in D(N_c, T_c), \quad f_c = T_c^{-1}, \\ \mathcal{L}(m, c) &= \begin{cases} c'_1 = c_2, \\ (c_1)_k = 0, & k \notin \{1, \dots, N_c\}, \\ (m_1)_k = 0, & \text{for } |k| \geq f_c/2, \\ (m_2/m_1)_k = 0, & \text{for } |k| \geq f_c/2. \end{cases} \end{aligned} \quad (5)$$

2.2 Separation problem

In the separation problem we are interested to solve, there are two different situations to consider. First, the ideal situation described in Section 2.1, referred here as the *noise-free* issue. In real signal application, noise has to be taken into consideration, which is why an *optimal demodulation* has been proposed.

In the Fourier domain, the *Discrete Fourier Transform* (DFT) of the obtained discrete signal of total length N ($s(t_n)_{n=1, \dots, N}$), denoted here with a tilde sign \tilde{s} , namely,

$$\forall k = 1, \dots, N_c + N_m, \quad \tilde{s}[k] = \sum_{n=1}^N s(t_n) e^{-i2\pi \frac{k}{N} n}, \quad (6)$$

is then a circular convolution of the carrier and modulation DFTs, i.e., $\tilde{s} = \sum_{j=1}^q \tilde{c}_j * \tilde{m}_j$.

In the case where the patterns centered on two consecutive harmonics do not overlap, the spectrum of the modulation is directly accessible from the spectrum of \tilde{s} . This is the well-known condition making demodulation possible.

Hypothesis 1. In what follows, we shall assume that:

$$2N_m f_m < f_c. \quad (7)$$

Note that $(c_j(t_n))_{n=1, \dots, N}$ is T_c -periodic with $T_c = N_c T_s$, i.e., $c_j(t_{n+N_c}) = c_j(t_n)$, $(m_j(t_n))_{n=1, \dots, N}$ is T_m -periodic with $T_m = N_m T_s$, i.e., $m_j(t_{n+N_m}) = m_j(t_n)$.

In the noise-free situation, finding those couples will be referred as Problem 1.

Problem 1. (Exact demodulation). Given $N \in \mathbb{N}$ and

- (1) a discrete signal $(s(t_n))_{n=1, \dots, N}$ of sampling period T_s , i.e., $t_n = n T_s$, and a duration $T_{tot} = N T_s$,

- (2) a frequency $f_c = k_c \cdot f_{tot}$, where $f_{tot} = 1/T_{tot}$ and $k_c \in \mathbb{N}$ is such that $N = k_c N_c$ for a certain $N_c \in \mathbb{N}$,

the problem aims to find a sequence of carrier/modulation pairs $((c_j(t_n), m_j(t_n)))_{n=1, \dots, N}$ satisfying

$$\forall n = 1, \dots, N, \quad s(t_n) = \sum_{j=1}^q c_j(t_n) m_j(t_n),$$

where $(c_j(t_n))_{n=1, \dots, N}$ is T_c -periodic with $T_c = N_c T_s$, $(m_j(t_n))_{n=1, \dots, N}$ is T_{m_j} -periodic with $T_{m_j} = N_{m_j} T_s$, and $(c_j(t_n), m_j(t_n))$ verifying Hypothesis 1.

However, when we consider a more realistic signal, i.e., a noisy signal, we have to use an optimal demodulation. Formally, we seek for a sequence of pairs of signals $((\hat{c}_j, \hat{m}_j))_{j=1, \dots, q}$ solving the optimization problem.

Problem 2. (Optimal demodulation). Given (1) and (2) as stated in Problem 1, find a sequence carrier/modulation pairs $((c_j(t_n), m_j(t_n)))_{n=1, \dots, N}$ which minimizes the following cost function

$$C(c, m) = \sum_{n=1}^N \left| \sum_{j=1}^q (c_j(t_n) m_j(t_n)) - s(t_n) \right|^2$$

over all carrier/modulation pairs satisfying Hypothesis 1, and where $(c_j(t_n))_{n=1, \dots, N}$ is T_c -periodic and $(m_j(t_n))_{n=1, \dots, N}$ is T_{m_j} -periodic.

2.3 Matrix formulation

In practice, a common specific case allows a very handy reformulation of the problem. Let $(s(t_n))_{n=1, \dots, N}$ be a discretized time signal and \tilde{s} its DFT defined by (6). The matrix representation we propose consists in cutting \tilde{s} into buckets centered on each harmonic of the carrier and then stacking them vertically as illustrated by Figure 1. Let us

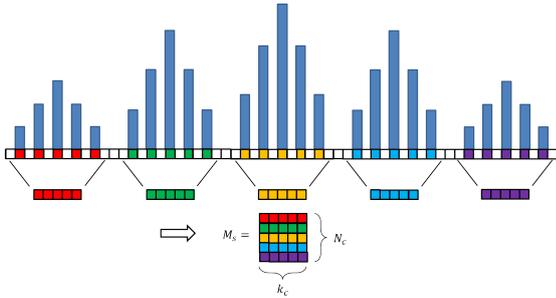


Figure 1. Matrix spectrum construction from the line spectrum vector

write down a mathematical definition of this matrix M_s .

Definition 1. (Matrix representation of a spectrum). Let $(s(t_n))_{n=1, \dots, N}$ be a discrete time signal of length N , \tilde{s} its DFT defined by (6), and k_c an integer dividing N . We call *matrix representation of \tilde{s} for k_c periods* the matrix

$$[M_s]_{i+1, :} = \tilde{s} [I_{k_c} + k_c i], \quad (8)$$

where $I_{k_c} + k_c i$ denotes the sequence I_{k_c} with $k_c i$ added to each element. This definition is illustrated by Figure 1.

Problems 1 and 2 can be both reformulated within a matrix formulation.

Problem 3. (Exact demodulation). If \tilde{c}_j (resp., \tilde{m}_j) denotes the vector containing the Fourier coefficients of $c_j(t_n)$ (resp., $m_j(t_n)$), and H the Hermitian transpose operator, then Problem 2 is equivalent to finding the vectors \tilde{c}_j 's and the vectors \tilde{m}_j 's satisfying:

$$M_s = \sum_{j=1}^q \tilde{c}_j \tilde{m}_j^H.$$

Problem 4. (Optimal demodulation). If \tilde{c}_j (resp., \tilde{m}_j) denotes the vector containing the Fourier coefficients of $c_j(t_n)$ (resp., $m_j(t_n)$), and H the Hermitian transpose operator, then Problem 2 is equivalent to finding the vectors \tilde{c}_j 's and the vectors \tilde{m}_j 's which minimizes the function

$$C(\tilde{c}, \tilde{m}) = \left\| \sum_{j=1}^q \tilde{c}_j \tilde{m}_j^H - M_s \right\|_{Fro}^2,$$

where $\|M\|_{Fro} = \left(\sum_{i=1}^n \sum_{j=1}^m |M_{ij}|^2 \right)^{1/2}$ denotes the so-called *Frobenius norm* of $M \in \mathbb{K}^{n \times m}$.

Due to length constraint of this paper, we shall only consider here Problem 3. For the study of Problem 4, we refer the reader to Hubert et al. (2019).

In the application case we are interested in, i.e., *phase and amplitude modulation* signals, Problem 3 gives a reformulation of (4) in the following matrix form

$$M_s = \tilde{c}_1 \tilde{m}_1^H + \tilde{c}_2 \tilde{m}_2^H,$$

where $\tilde{c}_2 = \tilde{c}_1'$ is the derivative of \tilde{c}_1 in the frequency domain. This derivative operation can be expressed as $\tilde{c}_1' = D \tilde{c}_1$, where D is the diagonal matrix computing the derivative in the frequency domain, i.e.:

$$D = i 2 \pi f_c \text{diag}(-N_c, \dots, N_c).$$

Thus, we obtain the following general formulation of the main problem considered in this paper:

$$M_s = \tilde{c}_1 \tilde{m}_1^H + D \tilde{c}_1 \tilde{m}_2^H.$$

3. SOLVING INHOMOGENEOUS LINEAR SYSTEMS

3.1 Notation & basic homological algebra

In what follows, \mathbb{K} will denote a field (e.g., $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$) and $A \in \mathbb{K}^{r \times s}$ a $r \times s$ matrix with entries in \mathbb{K} . Associated with $A \in \mathbb{K}^{r \times s}$, we can consider the two \mathbb{K} -linear maps:

$$\begin{aligned} \cdot A : \mathbb{K}^{1 \times r} &\longrightarrow \mathbb{K}^{1 \times s} & A \cdot : \mathbb{K}^{s \times 1} &\longrightarrow \mathbb{K}^{r \times 1} \\ \lambda &\longmapsto \lambda A, & \eta &\longmapsto A \eta. \end{aligned} \quad (9)$$

If $B \in \mathbb{K}^{s \times t}$ is such that $A B = 0$, then we can consider the so-called *complexes* of \mathbb{K} -vector spaces (Rotman (2009))

$$\mathbb{K}^{1 \times r} \xrightarrow{\cdot A} \mathbb{K}^{1 \times s} \xrightarrow{A \cdot} \mathbb{K}^{r \times 1}, \quad (10)$$

$$\mathbb{K}^{r \times 1} \xleftarrow{A \cdot} \mathbb{K}^{s \times 1} \xleftarrow{A \cdot} \mathbb{K}^{t \times 1}, \quad (11)$$

i.e., linear maps with zero consecutive compositions, i.e.:

$$\begin{aligned} \text{im}_{\mathbb{K}}(\cdot A) &:= \mathbb{K}^{1 \times r} A \subseteq \ker_{\mathbb{K}}(A \cdot) := \{\mu \in \mathbb{K}^{1 \times s} \mid \mu B = 0\}, \\ \text{im}_{\mathbb{K}}(A \cdot) &:= B \mathbb{K}^{t \times 1} \subseteq \ker_{\mathbb{K}}(\cdot A) := \{\eta \in \mathbb{K}^{s \times 1} \mid A \eta = 0\}. \end{aligned}$$

The *defect of exactness* of (10) (resp., (11)) is defined by

$$H(\mathbb{K}^{1 \times s}) := \ker_{\mathbb{K}}(A \cdot) / \text{im}_{\mathbb{K}}(\cdot A)$$

$$\text{(resp., } H(\mathbb{K}^{s \times 1}) := \ker_{\mathbb{K}}(\cdot A) / \text{im}_{\mathbb{K}}(A \cdot)),$$

where E/F stands for the *quotient* of a \mathbb{K} -vector space E by a \mathbb{K} -subvector space F . E/F is the \mathbb{K} -vector space

defined by the *residue classes* $\pi(e)$ for all $e \in E$, where $\pi(e_1) = \pi(e_2)$ if $e_1 - e_2 \in F$, endowed with the operations:

$$\forall e_1, e_2 \in E, \forall k \in \mathbb{K}, \begin{cases} \pi(e_1) + \pi(e_2) := \pi(e_1 + e_2), \\ k \pi(e_1) := \pi(k e_1). \end{cases}$$

The complex (10) (resp., (11)) is said to be *exact* at $\mathbb{K}^{1 \times s}$ (resp., $\mathbb{K}^{s \times 1}$) if $H(\mathbb{K}^{1 \times s}) = 0$ (resp., $H(\mathbb{K}^{s \times 1}) = 0$), i.e., if $\ker_{\mathbb{K}}(.B) = \text{im}_{\mathbb{K}}(.A)$ (resp., $\ker_{\mathbb{K}}(A.) = \text{im}_{\mathbb{K}}(B.)$).

An exact sequence of the form $0 \longrightarrow \mathbb{K}^{1 \times s} \xrightarrow{.B} \mathbb{K}^{1 \times t}$ means that $\ker_{\mathbb{K}}(.B) = 0$, i.e., that $.B$ is injective, whereas the exact sequence $\mathbb{K}^{1 \times r} \xrightarrow{.A} \mathbb{K}^{1 \times s} \longrightarrow 0$ means that $\text{im}_{\mathbb{K}}(.A) = \ker_{\mathbb{K}}(0) = \mathbb{K}^{1 \times s}$, i.e., that $.A$ is surjective. Similar comments hold for linear maps of the form $(A.)$.

A standard result on the duality of \mathbb{K} -vector spaces and \mathbb{K} -linear maps asserts that if (10) (resp., (11)) is an exact sequence of finite-dimensional \mathbb{K} -vector spaces, then so is (11) (resp., (10)). In homological algebra, we say that the *functor* $\text{hom}_{\mathbb{K}}(\cdot, \mathbb{K})$ is *exact*, where $\text{hom}_{\mathbb{K}}(E, \mathbb{K})$ denotes the \mathbb{K} -vector space formed by all the \mathbb{K} -linear maps (forms) from a \mathbb{K} -vector space E to \mathbb{K} . See, e.g., Rotman (2009).

Similarly a standard result in linear algebra asserts that if (10) (resp., (11)) is an exact sequence of finite-dimensional \mathbb{K} -vector spaces, then, for any $q \in \mathbb{N}$, so is

$$\mathbb{K}^{q \times r} \xrightarrow{.A} \mathbb{K}^{q \times s} \xrightarrow{.B} \mathbb{K}^{q \times t}, \quad (12)$$

$$\left(\text{resp., } \mathbb{K}^{r \times q} \xleftarrow{.A} \mathbb{K}^{s \times q} \xleftarrow{.B} \mathbb{K}^{t \times q} \right), \quad (13)$$

where $.A : \mathbb{K}^{q \times r} \longrightarrow \mathbb{K}^{q \times s}$ denotes the natural extension of (9) defined by $(.A)(\Lambda) := \Lambda A$ for all $\Lambda \in \mathbb{K}^{q \times r}$, and similarly for $.B$, $A.$ and $B.$ respectively. In homological algebra, we say that the *tensor functor* $\mathbb{K}^{q \times 1} \otimes_{\mathbb{K}} \cdot$ (resp., $\cdot \otimes_{\mathbb{K}} \mathbb{K}^{1 \times q}$) is exact, where $E \otimes_{\mathbb{K}} F$ stands for the *tensor product* of two finite-dimensional \mathbb{K} -vector spaces E and F . For more details, see, e.g., Rotman (2009).

If $0 \longrightarrow \mathbb{K}^{1 \times r} \xrightarrow{.A} \mathbb{K}^{1 \times s} \xrightarrow{.B} \mathbb{K}^{1 \times t} \longrightarrow 0$ is a short exact sequence of finite-dimensional \mathbb{K} -vector spaces, then the *Euler-Poincaré characteristic* asserts that $t - s + r = 0$. More generally, the Euler-Poincaré characteristic of a long exact sequence of finite-dimensional \mathbb{K} -vector spaces

$$0 \longrightarrow \mathbb{K}^{1 \times r_n} \xrightarrow{.A_n} \mathbb{K}^{1 \times r_{n-1}} \xrightarrow{.A_{n-1}} \dots \xrightarrow{.A_1} \mathbb{K}^{1 \times r_0} \longrightarrow 0$$

asserts that $\sum_{i=0}^n (-1)^i r_i = 0$. See e.g., Rotman (2009). Now, since the duality, i.e., the functor $\text{hom}_{\mathbb{K}}(\cdot, \mathbb{K})$, preserves the exactness of long exact sequences of finite-dimensional \mathbb{K} -vector spaces, the same result holds for a long exact sequence of the form:

$$0 \longleftarrow \mathbb{K}^{r_0 \times 1} \xleftarrow{.A_1} \mathbb{K}^{r_1 \times 1} \xleftarrow{.A_2} \dots \xleftarrow{.A_n} \mathbb{K}^{r_n \times 1} \longleftarrow 0.$$

Such an exact sequence always *splits* (Rotman (2009)), i.e., there exist $B_i \in \mathbb{K}^{r_{i-1} \times r_i}$ for $i = 0, \dots, n$, such that:

$$B_0 = B_{n+1} = 0, B_i A_i + A_{i+1} B_{i+1} = I_{r_i}, B_{i+1} B_i = 0.$$

3.2 A standard result of linear algebra

Let \mathbb{K} be field (e.g. $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$), $A \in \mathbb{K}^{r \times s}$ a $r \times s$ matrix with entries in \mathbb{K} , and $y \in \mathbb{K}^{r \times t}$. We first state a standard result on the existence of solutions $x \in \mathbb{K}^{s \times t}$ of the following \mathbb{K} -linear inhomogeneous system:

$$A x = y. \quad (14)$$

Let us consider the following \mathbb{K} -vector space:

$$\ker_{\mathbb{K}}(.A) := \{\lambda \in \mathbb{K}^{1 \times r} \mid \lambda A = 0\}.$$

Let $B \in \mathbb{K}^{q \times r}$ be a matrix whose rows form a basis of $\ker_{\mathbb{K}}(.A)$. In other words, B is a *full row rank matrix* (i.e., $\mu B = 0$ yields $\mu = 0$ since the rows of B are \mathbb{K} -linearly independent) which satisfies:

$$\ker_{\mathbb{K}}(.A) = \text{im}_{\mathbb{K}}(.B) := \mathbb{K}^{1 \times q} B.$$

In particular, we have $q = \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.A))$.

Then, we have $\lambda y = \lambda A x = 0$ for all $\lambda \in \ker_{\mathbb{K}}(.A)$. Hence, $B y = 0$ is a necessary condition for the solvability of (14). The next standard theorem shows that it is also sufficient.

Theorem 1. For a fixed $A \in \mathbb{K}^{r \times s}$ and a fixed $y \in \mathbb{K}^{r \times t}$, the system (14) is solvable, i.e., (14) admits a solution $x \in \mathbb{K}^{s \times t}$, iff the following compatibility condition holds:

$$B y = 0. \quad (15)$$

Then, all the solutions of (14) are given by

$$\forall z \in \mathbb{K}^{u \times t}, \quad x = E y + C z, \quad (16)$$

where $C \in \mathbb{K}^{s \times u}$ is a matrix whose columns form a basis of $\ker_{\mathbb{K}}(A.) := \{\eta \in \mathbb{K}^{s \times 1} \mid A \eta = 0\}$, i.e., C is a *full column rank matrix* (i.e., $C \theta = 0$ yields $\theta = 0$) which satisfies

$$\ker_{\mathbb{K}}(A.) = \text{im}_{\mathbb{K}}(C.) := C \mathbb{K}^{u \times 1},$$

and $E \in \mathbb{K}^{s \times r}$ is a *generalized inverse* of A , i.e. $A E A = A$.

4. SOLUTION OF THE EXACT PROBLEM

4.1 Reformulation of the problem and a few remarks

Motivated by Problem 3, the main problem can be stated.

Rank 2 decomposition problem: Given $M \in \mathbb{K}^{n \times m}$ and $D \in \mathbb{K}^{n \times n}$, determine – if they exist – $u \in \mathbb{K}^{n \times 1}$ and $v_1, v_2 \in \mathbb{K}^{1 \times m}$ satisfying:

$$M = u v_1 + D u v_2. \quad (17)$$

Problem 3 corresponds to the case $\mathbb{K} = \mathbb{C}$, $q = 2$, i.e., the phase and amplitude modulation problem, $u = \tilde{c}_1$, $D u = \tilde{c}_2$, $v_1 = \tilde{m}_1^H$, and $v_2 = \tilde{m}_2^H$.

Let us note:

$$A(u) := (u \quad D u) \in \mathbb{K}^{n \times 2}, \quad v := (v_1^T \quad v_2^T)^T \in \mathbb{K}^{2 \times m}.$$

Then, (17) can be rewritten as:

$$A(u) v = M. \quad (18)$$

Problem defined in (18) is bilinear in u and v .

Remark 2. If $M = 0$, then $u = 0$ or $v_1 = v_2 = 0$ solves the problem. Hence, in what follows, we suppose that $M \neq 0$.

Remark 3. The \mathbb{K} -vector space $\text{im}_{\mathbb{K}}(A(u.))$ is generated by the two vectors u and $D u$, which shows that:

$$\text{rank}_{\mathbb{K}}(A(u)) := \dim_{\mathbb{K}}(\text{im}_{\mathbb{K}}(A(u.))) \leq 2.$$

A necessary condition for the solvability of (17) is then:

$$\text{rank}_{\mathbb{K}}(M) \leq 2. \quad (19)$$

Remark 4. If (18) is solvable with a non full row rank matrix v , then there exists $\alpha := (\alpha_1 \quad \alpha_2) \in \mathbb{K}^{1 \times 2}$ such that $\alpha v = \alpha_1 v_1 + \alpha_2 v_2 = 0$, which yields:

$$\begin{cases} M = ((D - \alpha_2 \alpha_1^{-1} I_n) u) v_2, & \text{if } \alpha_1 \neq 0, \\ M = ((I_n - \alpha_1 \alpha_2^{-1} D) u) v_1, & \text{if } \alpha_2 \neq 0. \end{cases}$$

Hence, $M \neq 0$ must satisfy $\text{rank}_{\mathbb{K}}(M) = 1$. A necessary condition for the solvability of (18) for a matrix M satisfying $\text{rank}_{\mathbb{K}}(M) = 2$ is then that v has full row rank.

Hence, we shall only consider full row rank matrix v .

4.2 Necessary condition for the solvability of Problem (17)

Let us solve (18). Let $L \in \mathbb{K}^{p \times n}$ be a matrix whose rows generate a basis of the \mathbb{K} -vector space $\ker_{\mathbb{K}}(.M)$, i.e., L is a full row rank matrix satisfying:

$$\ker_{\mathbb{K}}(.M) = \text{im}_{\mathbb{K}}(.L) = \mathbb{K}^{1 \times p} L.$$

Let us explicitly characterize p in terms of the rank of M :

$$\begin{aligned} p &= \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.M)) = n - \dim_{\mathbb{K}}(\text{im}_{\mathbb{K}}(.M)) \\ &= n - \dim_{\mathbb{K}}(\text{im}_{\mathbb{K}}(M.)) = n - \text{rank}_{\mathbb{K}}(M). \end{aligned} \quad (20)$$

Now, (18) yields:

$$L A(u) v = L M = 0. \quad (21)$$

Since v has full row rank, we get $L A(u) = 0$, i.e., u must satisfy the following \mathbb{K} -linear system:

$$\begin{cases} L u = 0, \\ L D u = 0. \end{cases} \quad (22)$$

Remark 5. Since the p rows of L are \mathbb{K} -linearly independent, the dimension of the \mathbb{K} -vector solution space of $L u = 0$, i.e., $\ker_{\mathbb{K}}(L.)$, is $n - p = \text{rank}_{\mathbb{K}}(M)$ by (20). The dimension of the solution space of (22) is then at most $\text{rank}_{\mathbb{K}}(M)$ (exactly $\text{rank}_{\mathbb{K}}(M)$ if, e.g., $D = I_n$ or $L D = 0$).

Let us now derive an equivalent characterization of (22). By definition of L , we have the following exact sequence:

$$0 \longrightarrow \mathbb{K}^{1 \times p} \xrightarrow{.L} \mathbb{K}^{1 \times n} \xrightarrow{.M} \mathbb{K}^{1 \times m}.$$

Applying the exact functor $\text{hom}_{\mathbb{K}}(\cdot, \mathbb{K})$ to it, we obtain the following dual exact sequence of \mathbb{K} -vector spaces:

$$0 \longleftarrow \mathbb{K}^{p \times 1} \xleftarrow{L.} \mathbb{K}^{n \times 1} \xleftarrow{M.} \mathbb{K}^{m \times 1}.$$

Using $\ker_{\mathbb{K}}(L.) = \text{im}_{\mathbb{K}}(M.)$, we get $L u = 0$ is equivalent to the existence of $w \in \mathbb{K}^{m \times 1}$ such that $u = M w$. Thus, the second equation of (22) is equivalent to $L(D M w) = 0$, which in turn is equivalent to the existence of $w' \in \mathbb{K}^{m \times 1}$ such that $D M w = M w'$, which can be rewritten as:

$$(M \quad - D M) \begin{pmatrix} w' \\ w \end{pmatrix} = 0.$$

Using (19) and the upper bound of *Sylvester's inequality*

$$\text{rank}_{\mathbb{K}}(D M) \leq \min\{\text{rank}_{\mathbb{K}}(D), \text{rank}_{\mathbb{K}}(M)\},$$

then (22) is equivalent to:

$$\text{rank}_{\mathbb{K}}(M \quad - D M) \leq 3. \quad (23)$$

Lemma 6. With the above notations, a necessary condition on u for the existence of a solution of Problem (17) is (22) with $p \geq n - 2$, or equivalently (19) and (23).

4.3 Necessary and sufficient conditions

Let u be a non-trivial solution of (22). We can now form the matrix $A(u) = (u \quad D u)$ and we are then led to the study of the linear inhomogeneous system $A(u) v = M$. Let $L' \in \mathbb{K}^{p' \times n}$ be a full row rank matrix whose rows form a basis of $\ker_{\mathbb{K}}(.A(u))$, i.e., $\ker_{\mathbb{K}}(.A(u)) = \text{im}_{\mathbb{K}}(.L')$, and $p' = \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.A(u)))$. By Theorem 1, there exists $v \in \mathbb{K}^{2 \times m}$ which satisfies $A(u) v = M$ iff the following compatibility condition holds:

$$L' M = 0. \quad (24)$$

Notice that the compatibility condition (24) depends on u , and thus we seek for u – if it exists – in the solution space of (22) so that (24) holds.

Let us reinterpret (24) and (21) in a more intrinsic mathematical setting. By definition of L , we have the following exact sequence of \mathbb{K} -vector spaces:

$$0 \longrightarrow \mathbb{K}^{1 \times p} \xrightarrow{.L} \mathbb{K}^{1 \times n} \xrightarrow{.M} \mathbb{K}^{1 \times m}.$$

Then, $L A(u) = 0$ iff u satisfies (22). If so, then we get the following complex of \mathbb{K} -vector spaces:

$$0 \longrightarrow \mathbb{K}^{1 \times p} \xrightarrow{.L} \mathbb{K}^{1 \times n} \xrightarrow{.A(u)} \mathbb{K}^{1 \times 2}.$$

The defect of exactness of this complex at $\mathbb{K}^{1 \times n}$ is then:

$$H(\mathbb{K}^{1 \times n}, u) := \ker_{\mathbb{K}}(.A(u)) / \text{im}_{\mathbb{K}}(.L).$$

Now, by definition of L' , we have the following exact sequence of \mathbb{K} -vector spaces:

$$0 \longrightarrow \mathbb{K}^{1 \times p'} \xrightarrow{.L'} \mathbb{K}^{1 \times n} \xrightarrow{.A(u)} \mathbb{K}^{1 \times 2}. \quad (25)$$

Hence, we obtain:

$$\begin{aligned} H(\mathbb{K}^{1 \times n}, u) &= \ker_{\mathbb{K}}(.A(u)) / \ker_{\mathbb{K}}(.M) \\ &= \text{im}_{\mathbb{K}}(.L') / \text{im}_{\mathbb{K}}(.L). \end{aligned}$$

Since $\text{im}_{\mathbb{K}}(.L) \subseteq \text{im}_{\mathbb{K}}(.L')$, $L \in \text{im}_{\mathbb{K}}(.L')$, and thus there exists $L'' \in \mathbb{K}^{p' \times p'}$ such that $L = L'' L'$. Since L' has full row rank, we obtain the following isomorphism:

$$H(\mathbb{K}^{1 \times n}, u) = \text{im}_{\mathbb{K}}(.L') / \text{im}_{\mathbb{K}}(.L) \cong \mathbb{K}^{1 \times p'} / \text{im}_{\mathbb{K}}(.L'').$$

Hence, $H(\mathbb{K}^{1 \times n}, u) = 0$ iff $\text{im}_{\mathbb{K}}(.L'') = \mathbb{K}^{1 \times p'}$, i.e., iff there exists $X \in \mathbb{K}^{p' \times p'}$ such that $X L'' = I_{p'}$, i.e., iff L'' admits a left inverse, which is also equivalent to the injectivity of the \mathbb{K} -linear map $L'' : \mathbb{K}^{p' \times 1} \longrightarrow \mathbb{K}^{p' \times 1}$.

Applying the exact functor $\text{hom}_{\mathbb{K}}(\cdot, \mathbb{K})$ to the exact sequence (25), we obtain the exact sequence:

$$0 \longleftarrow \mathbb{K}^{p \times 1} \xleftarrow{L'.} \mathbb{K}^{n \times 1} \xleftarrow{A(u).} \mathbb{K}^{2 \times 1}.$$

Then, applying the exact functor $\cdot \otimes_{\mathbb{K}} \mathbb{K}^{1 \times m}$ to the last exact sequence, we get the following exact sequence:

$$0 \longleftarrow \mathbb{K}^{p \times m} \xleftarrow{L'.} \mathbb{K}^{n \times m} \xleftarrow{A(u).} \mathbb{K}^{2 \times m}.$$

Hence, M belongs to $\text{im}_{\mathbb{K}^{1 \times m}}(A(u).) = A(u) \mathbb{K}^{2 \times m}$, i.e., there exists $v \in \mathbb{K}^{2 \times m}$ such that $M = A(u) v$, iff $L' M = 0$. By definition of L , we have $L M = 0$, i.e., $L'' (L' M) = 0$. Therefore, $L M = 0$ yields $L' M = 0$ iff L'' is injective, i.e., iff L'' admits a left inverse, i.e., iff $H(\mathbb{K}^{1 \times n}, u) \cong \mathbb{K}^{1 \times p'} / \text{im}_{\mathbb{K}}(.L'') = 0$, i.e., iff:

$$\ker_{\mathbb{K}}(.A(u)) = \ker_{\mathbb{K}}(.M). \quad (26)$$

Problem (17) is reduced to finding $0 \neq u$ satisfying (22) such that the \mathbb{K} -vector space $H(\mathbb{K}^{1 \times n}, u)$ is trivial, i.e.:

$$\dim_{\mathbb{K}}(H(\mathbb{K}^{1 \times n}, u)) = 0.$$

Remark 7. Let us give a direct interpretation of (26). Let us first suppose that (17) or equivalently that (18) is solvable and let us compare the \mathbb{K} -vector spaces:

$$\begin{aligned} \ker_{\mathbb{K}}(.A(u)) &= \{\lambda \in \mathbb{K}^{1 \times n} \mid \lambda A(u) = 0\}, \\ \ker_{\mathbb{K}}(.M) &= \{\lambda \in \mathbb{K}^{1 \times n} \mid \lambda M = 0\}. \end{aligned}$$

We clearly have $\ker_{\mathbb{K}}(.A(u)) \subseteq \ker_{\mathbb{K}}(.M)$. Now, if we consider $\lambda \in \ker_{\mathbb{K}}(.M)$, then $\lambda A(u) v = \lambda M = 0$, which yields $\lambda A(u) = 0$ since v is full rank. Thus, if a solution exists for Problem (17), then (26) holds.

Conversely, if u is a non-trivial solution of (22) such that (26), then $\mathbb{K}^{1 \times p'} L' = \mathbb{K}^{1 \times p} L$, which shows that $p' = p$ and there exist $U, V \in \mathbb{K}^{p \times p}$ such that $L' = U L$ and $L = V L'$, which yields $(U V - I_p) L' = 0$ and $(V U - I_p) L = 0$, i.e., $U V = I_p$ and $V U = I_p$ since both L and L' have full row rank. The compatibility condition $L' M = 0$ is thus equivalent to $L M = 0$, which is satisfied by definition of L . Then, Theorem 1 shows that there exists $v \in \mathbb{K}^{2 \times m}$ such that $A(u) v = M$, which solves (17).

Let us state the first main result of this paper.

Theorem 8. With the above notations, Problem (17) is solvable iff there exists $0 \neq u \in \mathbb{K}^{n \times 1}$ satisfying (22) and such that $H(\mathbb{K}^{1 \times n}, u) = \ker_{\mathbb{K}}(.A(u)) / \ker_{\mathbb{K}}(.M) = 0$.

Let us study the \mathbb{K} -vector space $H(\mathbb{K}^{1 \times n}, u)$. The Euler-Poincaré characteristic of the short exact sequence

$$0 \longrightarrow \ker_{\mathbb{K}}(.M) \xrightarrow{i} \ker_{\mathbb{K}}(.A(u)) \xrightarrow{\pi} H(\mathbb{K}^{1 \times n}, u) \longrightarrow 0$$

yields:

$$\dim_{\mathbb{K}}(H(\mathbb{K}^{1 \times n}, u)) = \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.A(u)) - \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.M))) = p' - p. \quad (27)$$

Let us now characterize p' . Considering the following two short exact sequences of \mathbb{K} -vector spaces

$$0 \longrightarrow \ker_{\mathbb{K}}(.A(u)) \longrightarrow \mathbb{K}^{1 \times n} \longrightarrow \text{im}_{\mathbb{K}}(.A(u)) \longrightarrow 0,$$

$$0 \longleftarrow \text{im}_{\mathbb{K}}(A(u)) \longleftarrow \mathbb{K}^{2 \times 1} \longleftarrow \ker_{\mathbb{K}}(A(u)) \longleftarrow 0,$$

the Euler-Poincaré characteristic then yields:

$$\begin{cases} \dim_{\mathbb{K}}(\text{im}_{\mathbb{K}}(.A(u))) = n - \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.A(u))) = n - p', \\ \dim_{\mathbb{K}}(\text{im}_{\mathbb{K}}(A(u))) = 2 - \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(A(u))). \end{cases}$$

Thus, we obtain $p' = n - 2 + \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(A(u)))$. By (20), we have $p = n - \text{rank}_{\mathbb{K}}(M)$. Hence, we obtain:

$$H(\mathbb{K}^{1 \times n}, u) = 0 \Leftrightarrow p' = p \\ \Leftrightarrow \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(A(u))) = 2 - \text{rank}_{\mathbb{K}}(M).$$

Corollary 9. With the above notations, Problem (17) is solvable iff there exists $0 \neq u \in \mathbb{K}^{n \times 1}$ satisfying (22) and one of the following two equivalent conditions holds:

- (1) $p' = p$, i.e., $\dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.A(u))) = \dim_{\mathbb{K}}(\ker_{\mathbb{K}}(.M))$,
- (2) $\dim_{\mathbb{K}}(\ker_{\mathbb{K}}(A(u))) = 2 - \text{rank}_{\mathbb{K}}(M)$.

Let us now study $\ker_{\mathbb{K}}(A(u)) = \{w \in \mathbb{K}^{2 \times 1} \mid A(u) w = 0\}$. If $w = (w_1 \ w_2) \in \ker_{\mathbb{K}}(A(u))$, i.e., $u w_1 + D u w_2 = 0$, then, using $u \neq 0$, we have $w = 0$ if $w_2 = 0$, or $D u = -w_1 w_2^{-1} u$ if $w_2 \neq 0$, i.e., u is an eigenvector of D with the eigenvalue $-w_1 w_2^{-1} \in \mathbb{K}$. Hence, if u is not an eigenvector of D , then $\ker_{\mathbb{K}}(A(u)) = 0$.

Now, if u is an eigenvalue of D with eigenvalue $\lambda \in \mathbb{K}$, then $\ker_{\mathbb{K}}(A(u)) = \text{im}_{\mathbb{K}}(K)$, where $K = (-\lambda \ 1)^T$, is a \mathbb{K} -vector space of dimension 1.

Corollary 10. With the above notations, Problem (17) is solvable iff there exists $0 \neq u \in \mathbb{K}^{n \times 1}$ satisfying (22) and:

- (1) If $\text{rank}_{\mathbb{K}}(M) = 2$, then u is not an eigenvector of D for an eigenvalue $\lambda \in \mathbb{K}$. Then, there exists a unique $v \in \mathbb{K}^{2 \times m}$ satisfying $A(u) v = M$ defined by

$$v = E M,$$

where $E \in \mathbb{K}^{2 \times n}$ denotes a left inverse of $A(u)$.

- (2) If $\text{rank}_{\mathbb{K}}(M) = 1$, then u is an eigenvector of D with an eigenvalue $\lambda \in \mathbb{K}$. Then, all the solutions $v \in \mathbb{K}^{2 \times m}$ satisfying $A(u) v = M$ are defined by

$$\forall z \in \mathbb{K}^{1 \times m}, \quad v = E M + K z,$$

where $K = (-\lambda \ 1)^T$ and $E \in \mathbb{K}^{2 \times n}$ denotes a generalized inverse of $A(u)$.

Remark 11. Let M satisfy $\text{rank}_{\mathbb{K}}(M) = 1$ and $D = 0$. Then, (20) shows that $p = n - 1$, and thus there exists a non-trivial solution u of the linear system $L u = 0$ when $n \geq 2$. Hence, 2 of Corollary 10 shows that M can always be written as $M = u v_1$, where $u \in \mathbb{K}^{n \times 1}$ and $v_1 \in \mathbb{K}^{1 \times m}$.

5. CONCLUSION

This paper develops a general class of problems originating from gearbox vibration analysis. Based on Hubert et al. (2018) that made a link between the fields of signal demodulation and low-rank approximation, in this paper, we studied the specific case of phase and amplitude demodulation. Rewritten it as a polynomial problem, a solution has been proposed for the exact problem along with the necessary and sufficient conditions for its solvability. In future work, the optimal problem will be studied as well as generalizations stated in a more general framework.

REFERENCES

- M. Feldman. Hilbert transform in vibration analysis. *Mechanical systems and signal processing*, 25(3):735–802, 2011.
- Z. Feng, F. Chu, and M. J. Zuo. Time-frequency analysis of time-varying modulated signals based on improved energy separation by iterative generalized demodulation. *Journal of Sound and Vibration*, 330(6):1225–1243, 2011.
- B. Friedlander and J. M. Francos. Estimation of amplitude and phase parameters of multicomponent signals. *IEEE Transactions on Signal Processing*, 43(4):917–926, 1995.
- N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: Mathematical, physical and engineering sciences*, volume 454, pages 903–995. The Royal Society, 1998.
- E. Hubert, A. Barrau, and M. El Badaoui. New multi-carrier demodulation method applied to gearbox vibration analysis. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2141–2145. IEEE, 2018.
- E. Hubert, A. Barrau, Y. Bouzidi, R. Dagher, and A. Quadrat. Algebraic aspects of the optimal demodulation problem. In preparation, 2019.
- J. F. Kaiser. On a simple algorithm to calculate the energy of a signal. In *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, pages 381–384. IEEE, 1990.
- P. D. McFadden. Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration. *Journal of vibration, acoustics, stress, and reliability in design*, 108(2):165–170, 1986.
- A. Potamianos and P. Maragos. A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation. *Signal Processing*, 37(1):95–120, 1994.
- J. J. Rotman. *Introduction to Homological Algebra*. Springer, 2009.
- B. Santhanam and P. Maragos. Multicomponent AM-FM demodulation via periodicity-based algebraic separation and energy-based demodulation. *IEEE Transactions on Communications*, 48(3):473–490, 2000.
- Z. Yu, Y. Sun, and W. Jin. A novel generalized demodulation approach for multi-component signals. *Signal Processing*, 118:188–202, 2016.