UNIVERSITÀ DEGLI STUDI DI UDINE

Dipartimento Politecnico di Ingegneria e Architettura

Ph.D. in Industrial and Information Engineering

XXXVIII Cycle

OPTICAL METHODS FOR VIBRO-ACOUSTIC MEASUREMENTS

Ph.D. Student

Dott.ssa Sofia Baldini

Supervisor

Prof. Paolo Gardonio

Year 2026

To those who know how to listen.

ABSTRACT

The accurate characterization of structural vibrations and their associated acoustic radiation represent a key task in vibro-acoustic research, with direct implications for design, diagnostics, and control. Conventional measurement approaches rely on accelerometers, laser vibrometers, microphone arrays, which, while accurate, are often costly, intrusive, and impractical for in situ applications. In this doctoral research, novel vision-based methodologies are investigated and developed as alternative routes to reconstruct vibration fields and sound radiation fields, with particular emphasis on total sound power radiation. Three case studies are presented, addressing both laboratory structures with simple geometries and more realistic structures as closed shells. The potential of emerging event-based cameras is also explored with the intent to develop a numerically-lightweight vibration image processing. Overall, the objective is to advance the feasibility, accuracy, and applicability of vision-based vibro-acoustic measurements. The first contribution concerns the optical reconstruction of vibration and acoustic radiation from a thin-walled cylindrical shell. A novel framework is introduced that integrates high-speed camera acquisition with digital image correlation and photogrammetry to recover the full displacement field of the structure. The reconstructed vibration fields are validated against reference laser vibrometer data, while the radiated acoustic fields are benchmarked against a microphone array. Deviations are confined within $1-3\,\mathrm{dB}$ across narrow-band and third-octave-band spectra. A general formulation is proposed in which the total radiated sound power is derived from surface-integrated sound intensity, approximated by a Riemann summation over elemental radiators. This approach demonstrates that camera-based methods can achieve accuracies comparable to traditional acoustic measurements with microphones or sound intensity probes, while requiring fewer measurement sensors (the camera or cameras) and offering non-intrusive, full-field acquisition. The second contribution extends the methodology to a machine-like vibro-acoustic source, a thin-walled metallic box, thereby shifting from laboratory model structures to a realistic application scenario. The study confirms the suitability of vision-based techniques for structures exhibiting discontinuities and geometric complexities. A further methodological advancement is introduced: the classical far-field integration of sound intensity prescribed by *ISO standards* is replaced by a novel formulation that collapses the Kirchhoff–Helmholtz integral onto the radiating surface itself. This formulation, implemented via the Boundary Element Method, allows the derivation of sound power directly from the reconstructed vibration fields, such that the sound radiation produced by complex geometries can be properly addressed. The results highlight both the accuracy and practical advantages of the proposed approach for in situ applications,

where conventional acoustic facilities may not be available. The third contribution explores the use of event-based cameras for vibration measurement. Unlike conventional frame-based cameras, event cameras asynchronously record pixel-level brightness changes, thus achieving microsecond-scale temporal resolution with minimal data redundancy. The full event processing is developed, encompassing event-to-frame reconstruction using the *E2VID* neural network, multi-camera calibration, subpixel marker tracking, and $3D$ triangulation. Experimental validation on a cantilever beam demonstrates accurate reconstruction of flexural vibration modes, with residual root mean square error below 4% with respect to the peak vibration amplitude when compared with laser vibrometer measurements. Parametric analysis with different camera configurations establishes the increased robustness of four synchronized cameras over two, while hardware synchronization ensures sub-microsecond temporal alignment. Beyond structural validation, the study highlights the potential of event cameras as compact, high-speed, and data-efficient sensors for vibro-acoustic applications, opening new perspectives for their use in capturing the vibration and acoustic radiation of complex machine systems. Overall, this doctoral work aims to consolidate and expand the role of vision-based methods in vibro-acoustics, showing that optical sensing, whether frame-based or event-driven, can provide dense, accurate, and non-intrusive measurements of vibration fields and radiated sound power. The proposed formulations and experimental validations produced in this thesis indicate that such techniques can match, and in some cases exceed, traditional vibro-acoustic measurement techniques in terms of measurement accuracy, while offering distinct advantages with respect to cost, scalability, flexibility, and applicability in real environments. The work presented in this thesis provides a foundation for future research at the intersection of structural dynamics, acoustics, and optical measurements, with promising implications for both laboratory studies and in situ monitoring of engineering structures and machines.

CONTENTS

1
INTRODUCTION

In this chapter, the background concepts and paradigms at the basis of the work presented in this thesis are introduced. It begins with a review of the fundaments of mechanical vibrations and sound radiation measurements, describing classical measurement methods and their typical limitations, especially for in situ testing. The chapter also outlines optical measurement techniques, with a focus on full-field vibration measurements using high-speed cameras [1–4]. In addition, the motivations for this research are outlined together with the main contributions of the thesis and its overall structure, providing the reader with the context needed to follow the methods and results discussed in the subsequent chapters.

1.1 FUNDAMENTALS OF VIBRATION AND ACOUSTIC MEASUREMENT CHAINS

Vibration denotes the study of oscillatory phenomena about equilibrium configurations, formulated through differential equations of motion derived from Newtonian, Lagrangian, or Hamiltonian principles, and encompassing linear and nonlinear, discrete and continuous systems. In this thesis, structural vibration is examined as an independent topic to characterize mechanical the dynamic responses of systems under various excitation conditions, while also establishing a rigorous basis for the derivation of acoustic fields induced by structural vibration. It is a multidisciplinary field spanning physics, engineering, applied mathematics, materials science and others. At its core, vibration analysis is governed by Newton's laws, expressed through differential equations of motion. For linear systems, the fundamental equations of motion are normally cast in the following matrix expression:

$$\boldsymbol{M}\ddot{\boldsymbol{x}}(t) + \boldsymbol{C}\dot{\boldsymbol{x}}(t) + \boldsymbol{K}\boldsymbol{x}(t) = \boldsymbol{F}(t), \tag{1.1}$$

where $\boldsymbol{x}(t)$ is the displacements vector (and its derivatives in time are respectively: $\dot{\boldsymbol{x}}(t)$, $\ddot{\boldsymbol{x}}(t)$), $\boldsymbol{M}$ represents the mass matrix of the system, $\boldsymbol{C}$ is the damping matrix, $\boldsymbol{K}$ is the stiffness matrix, and $\boldsymbol{F}(t)$ represents the external vector force acting on the system.

Solutions that satisfy both the governing equations of motion and the appropriate boundary conditions define the vibration field of the system, enabling prediction of its dynamic response under specified constraints. While vibration focuses on mechanical oscillations in solid structures, these vibrations can also propagate through surrounding media, leading to acoustic phenomena. Indeed, acoustics is an interdisciplinary field encompassing physics, engineering, signal processing, audiology, psychology, music, architecture, and other sciences. In this thesis, acoustics is approached as the study of mechanical vibrations and the propagation of the resulting wave phenomena through various media. Depending on frequency content and propagation medium, these vibrations can become audible and be perceived as sound. As a fundamentally mechanical phenomenon, the governing equations of acoustics can be derived directly from basic physical laws. These equations describe sound propagation in an unbounded medium. However, when accounting for interactions with physical structures, the governing equations must be complemented by boundary conditions specifying the constraints at interfaces. Solutions satisfying both the governing equations and boundary conditions define the sound field within a given system.

Basic linear acoustics deals with small-amplitude oscillations of pressure and particle velocity within the propagation medium. These oscillations are assumed to be small relative to the characteristic dimensions of the problem but large enough compared to molecular scales to justify a continuum description. Under this assumption, the medium is treated as continuous, and the fundamental equations of continuum mechanics remain valid. For a homogeneous fluid, the motion is governed by the conservation of mass, momentum, and energy. By combining the linearized continuity and momentum equations for an inviscid, homogeneous, isotropic fluid, the classical wave equation for acoustic pressure is obtained:

$$\frac{\partial^2 p(x_D, t)}{dt^2} = c^2 \nabla^2 p(x_D, t), \tag{1.2}$$

where c is the speed of sound in the medium and x_D is the position in it. This equation governs the propagation of small perturbations in pressure as acoustic waves and serves as the fundamental model for predicting sound fields in many practical engineering and scientific applications. Having established the governing equations, it is essential to consider how these phenomena are observed and quantified in practice, which naturally leads to the distinction between vibration and acoustic measurement. While both vibrations in solid structures and the resulting acoustic phenomena involve oscillatory motion, the methods for analysing and measuring them differ significantly. Vibration measurements are essential for characterizing the dynamic behaviour of structures, while acoustic measurements focus on capturing pressure fluctuations in fluids. In fact, vibration measurements play a crucial role not only during the design and development of engineering systems, but also for system characterization and operational monitoring. At various stages of a system lifecycle, such measurements may be required to provide experimental support for a structural dynamic model or to gain insight into the actual dynamic behaviour of the system under real conditions [5]. The difference between *vibration* and *sound* is typically defined by how we perceive or detect them, rather than by a strict physical or frequency-based separation. Vibration generally refers to the mechanical oscillatory motion of solid structures, often but not exclusively at relatively low frequencies, which can be perceived as whole-body motion or tactile sensations. Sound, on the other hand, is considered an audible phenomenon that is audible to the human ear and that primarily occurs in fluid media such as air. This distinction is therefore mainly practical and perceptual in nature and does not preclude the existence of high-frequency vibrations or of acoustic phenomena outside the human audible range, such as ultrasound. Nevertheless, vibrating surfaces are responsible for a significant share of noise generation across the entire audio-frequency spectrum, which is normally reported as *structure-borne noise*. Acoustic measurements are typically concerned with time-varying pressure fields in fluids. However, pressure is only one of several quantities that fluctuate over time in a medium supporting a sound wave; other variables include density, temperature, and the motion of fluid particles, specifically their position, velocity, and acceleration [6]. Among these, sound pressure is the most frequently measured parameter because it is the most accessible experimentally and, importantly, it is the direct physical stimulus for human hearing. The range of noise levels encountered in daily life can be extremely large. To manage this effectively, the sound pressure level (*SPL*) of continuous sounds is defined using a logarithmic scale. The primary audiological rationale for adopting a logarithmic metric lies in the fact that the human auditory system response to changes in sound pressure is not linear but approximately logarithmic. In both vibration and acoustic testing, a central quantity for characterizing system dynamics is the Frequency Response Function (*FRF*), which provides a unifying framework for relating input excitations to measurable system responses. The *FRF* describes the steady-state relationship between input and output in linear systems across a range of frequencies. Its specific form depends on the physical quantities involved (for example, force, acceleration, velocity, or displacement in structural testing, and pressure or volume velocity in acoustics). Table 1 summarises typical definitions of *FRFs* for structural applications [7]. Experimental determination of *FRFs* is a fundamental task in testing. These functions are commonly obtained through classical methods such as electrodynamic shaker and instrumented hammer testing for structures, or loudspeaker excitation and microphone measurements for acoustics. Once available, *FRFs* can be used not only to identify key modal parameters such as natural frequencies, mode shapes, and damping ratios, but also to support more advanced procedures like experimental modal analysis [8] or transfer path analysis (*TPA*) [7], particularly when operational loads are unknown. In practical applications, *FRFs* also serve a crucial role in assessing vibration isolation or noise control performance by characterising source, transmission, and receiver dynamics. Thus, the motivation for conducting vibration and acoustic measurements is rooted in the need to develop accurate dynamic models and understand how energy is transmitted through mechanical systems. *FRFs* serve as the foundational data product of these measurements and are indispensable for both design validation and diagnostic purposes.

Table 1: typical definitions of FRFs

Standard FRF	**Inverse FRF**
$receptance = \frac{displacement}{force}$	$dynamic\ stiffness = \frac{force}{displacement}$
$\times\, i\omega$	$\div\, i\omega$
$mobility = \frac{velocity}{force}$	$mechanical\ impedance = \frac{force}{velocity}$
$\times\, i\omega$	$\div\, i\omega$
$accelerance = \frac{acceleration}{force}$	$apparent\ mass = \frac{force}{acceleration}$
$trasmissibility = \frac{displacement}{displacement} = \frac{velocity}{velocity} = \frac{acceleration}{acceleration}$ or $= \frac{force}{force}$	

The typical measurement chain used for vibro-acoustic testing and the experimental acquisition of *FRFs* is composed of a series of elements, each playing a specific role in ensuring accurate and meaningful results. A schematic representation of this process is shown in Figure 1, which outlines the complete signal flow from excitation to data interpretation.

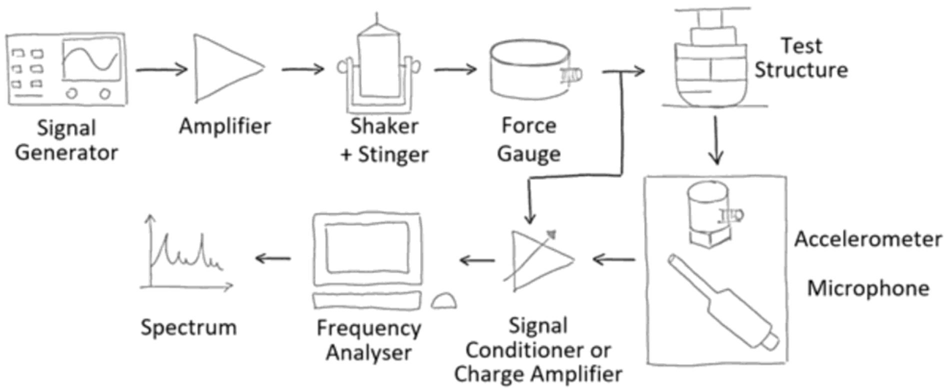


Figure 1: typical measurement chain used for vibro-acoustic testing

The process begins with the excitation of the structure, which can be achieved using an electrodynamic shaker or an instrumented impact hammer. In the case of shaker testing, the excitation signal, often generated directly by the measurement system, is amplified through a power amplifier and applied to the test structure via a mechanical coupling element, typically referred to as a *stinger*. This component ensures that the force is transmitted effectively while minimising interference with the system dynamics. In contrast, hammer excitation relies on a manual impact, with the force measured directly by a sensor embedded behind the hammer tip. Alternatively, in applications involving acoustic analysis, it is also common to analyse the sound pressure generated by the machine during its normal operation, without applying any artificial excitation. This approach allows the extraction of dynamic information directly from the operational sound field, and is especially relevant in the context of noise and vibration diagnostics or transfer path analysis. To capture the system response, a set of vibration or acoustic sensors is employed. The most commonly used are piezoelectric accelerometers, although laser vibrometers, strain gauges and microphones may be used depending on the specific requirements of the test, such as non-contact measurement, low-frequency sensitivity etc. The input force and output response signals are conditioned through appropriate electronics, typically charge amplifiers or signal conditioners, to make them compatible with the data acquisition system. The conditioned signals are then processed by a frequency analyser (or spectrum analyser), which performs the digitization of the analog signals and computes their Fourier transforms, usually via the Fast Fourier Transform (*FFT*) algorithm. This stage is critical for converting the time-domain signals into the frequency domain and for calculating the *FRF* as the ratio between output and input spectra. Finally, the measured *FRFs* are displayed and interpreted to extract dynamic characteristics of the structure, such as natural frequencies, mode shapes, and modal damping. These results form the basis for further analyses, including experimental modal analysis, vibration isolation assessment, and transfer path analysis.

1.2 CLASSICAL VIBRATION MEASUREMENT TECHNIQUES

To obtain the input-output quantities involved in the construction of *FRFs*, suitable vibration sensors are required. These quantities are only meaningful under the assumption of linear, steady-state system behaviour. These sensors measure mechanical quantities such as force, displacement, velocity, or acceleration. Within the frequency domain, these quantities are easily related by integration or differentiation, so it is generally sufficient to measure just one of displacement, velocity, or acceleration, in addition to force. This section provides a concise overview of commonly used vibration sensors, including piezoelectric accelerometers, force transducers, strain sensors, impedance heads, and laser vibrometers. For each of these devices, key aspects such as their working principles, limitations, mounting considerations are discussed.

FORCE GAUGES

Piezoelectric force sensors are the most common devices used to measure dynamic forces, exploiting the direct piezoelectric effect, whereby certain materials generate an electric charge in response to mechanical strain. Among these, quartz is frequently used due to its long-term stability and thermal robustness, while ceramic materials like *PZT* (lead zirconate titanate) are also widely adopted for their high sensitivity.

In a typical configuration, the sensing element, often a pair of quartz discs, is preloaded between two masses (as shown in Figure 2). When a dynamic force is applied, part of it induces a compressive strain on the piezoelectric element, generating a small electrical charge proportional to the applied force. The sensor sensitivity (e.g., in pC/N) allows for accurate calibration and force measurement.

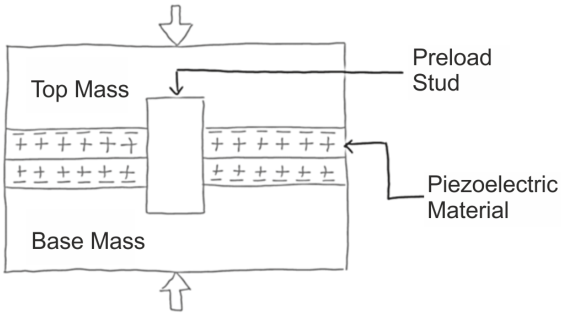


Figure 2: sketch of a force gauge piezoelectric sensor

ACCELEROMETERS

Accelerometers are fundamental sensors in structural dynamics and vibration testing, used to measure inertial acceleration of a point on a structure. They are available in various types based on different sensing principles, namely piezoelectric, piezoresistive, and variable capacitive, each suited to specific application domains. For frequency-domain measurements such as transfer function estimation, piezoelectric accelerometers are typically preferred due to their broad frequency bandwidth, which often spans from a few hertz up to tens of kilohertz.

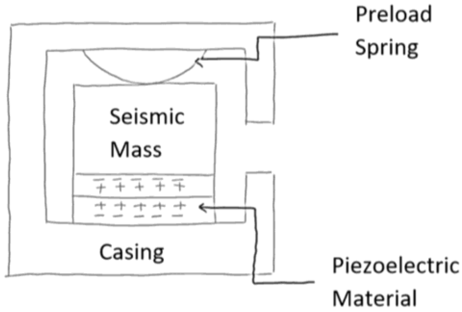


Figure 3: sketch of a piezoelectric accelerometer sensor

A schematic representation of a single-axis compression-type piezoelectric accelerometer is shown in Figure 3. It consists of a seismic mass m, mounted onto a sensing element composed of back-to-back piezoelectric discs held in compression by a preload spring, all enclosed within a rigid housing. When the base of the accelerometer is subjected to a dynamic input, inertial forces cause the seismic mass to accelerate relative to the base, resulting in mechanical deformation of the piezoelectric discs. This deformation generates an electrical output proportional to the acceleration.

The mechanical sensitivity of such a device can be analysed by modelling it as a single-degree-of-freedom (*SDOF*) system. Assuming harmonic base excitation of the form $Ae^{i\omega t}$, the frequency response function (*FRF*) between the relative displacement X of the mass and the base acceleration A is given by:

$$\frac{X}{A} = -\frac{1}{\omega_n^2}\left(\frac{1}{1-\left(\frac{\omega}{\omega_n}\right)^2 + 2\zeta i\frac{\omega}{\omega_n}}\right), \tag{1.3}$$

where ω is the excitation frequency, $\omega_n = \sqrt{k/m}$ is the natural frequency of the accelerometer, k is the effective stiffness of the sensing element and preload spring, ζ is the damping ratio and $i = \sqrt{-1}$ is the imaginary unit.

Considering frequencies which are much less than the natural frequency of the accelerometer, that is, $\omega/\omega_n \ll 1$, the mechanical sensitivity becomes approximately constant and independent from frequency: $X/A = 1/\omega_n^2$.

In general, larger accelerometers exhibit higher sensitivity due to their lower natural frequency but suffer from reduced usable bandwidth and may introduce mass loading effects on lightweight structures. As a practical rule, the measurement frequency range should remain below 20% of the sensor natural frequency to maintain a flat and accurate frequency response. However, this analysis assumes that the accelerometer is rigidly mounted to the structure, an idealization that rarely holds in practice. In reality, the interface through which the accelerometer is mounted introduces additional compliance, effectively forming a secondary dynamic system. These effects modify the sensor sensitivity according to the transmissibility across the interface and further restrict the usable frequency range of the measurement. The mounting method thus plays a critical role in shaping the overall frequency response of the system. To mitigate these issues, various mounting techniques are employed, each offering different trade-offs between mechanical stiffness, ease of installation, and bandwidth. Among them, stud mounting, where the sensor is attached via a threaded connection to a tapped hole, offers the highest contact stiffness and widest bandwidth, typically more than $10\ kHz$. When permanent modifications to the structure are not permissible, adhesive mounting techniques provide a viable alternative. For example, less rigid options such as wax or double-sided tape are acceptable for low-frequency and non-critical applications.

Additionally, magnetic mounts, though convenient for use on ferromagnetic surfaces, are generally restricted to frequencies below $1\ kHz$.

IMPEDANCE HEADS

Impedance heads are integrated sensors, as shown in Figure 4, which are composed by a force sensor and an accelerometer in series. This configuration allows for the direct estimation of the mechanical impedance at the driving point by analysing the ratio of force to acceleration signals in the frequency domain.

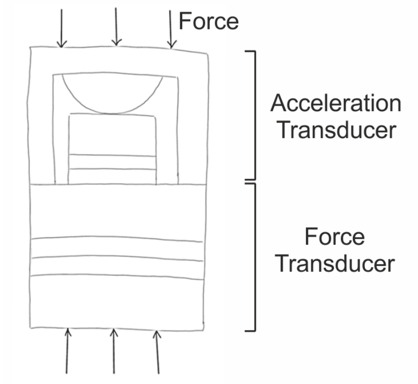


Figure 4: sketch of a piezoelectric impedance head sensor

STRAIN GAUGES

Strain sensors are employed to detect mechanical deformation and convert it into an electrical signal, enabling the quantification of strain within a structure or material. They are typically based on one of three transduction mechanisms: piezoresistive**,** piezoelectric, or capacitive. Among these, piezoresistive strain sensors are particularly widespread due to their mechanical robustness and straightforward signal conditioning. These sensors operate by correlating the applied mechanical load with strain-induced changes in electrical resistance which is typically measured through a Wheatstone bridge configuration. A typical strain gauge arranges a long, thin conductive strip in a zig-zag pattern of parallel lines as reported in Figure 5.

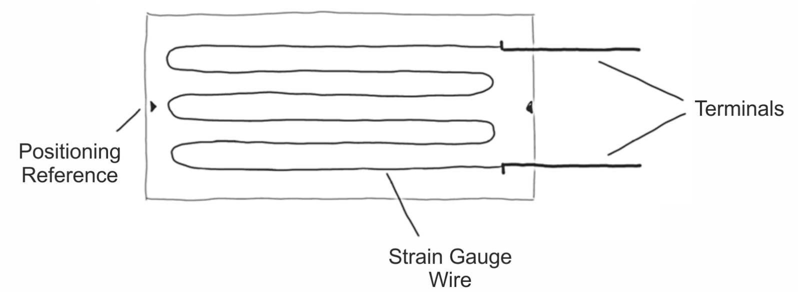


Figure 5: sketch of a strain gauge sensor

This configuration is typically implemented in the form of foil-type strain gauges, which are widely employed across numerous applications due to their versatility and relatively simple integration. In most cases, the orientation of the gauge with respect to the principal strain direction is a key factor for accurate measurement. The gauges are bonded to the substrate using specialized adhesives selected according to the intended duration of measurement.

LASER DOPPLER VIBROMETER

The primary limitations of contact-based measurement techniques involve the added mass and local stiffening they introduce to the structure, as well as their restriction to measurements at discrete points on the test object. To overcome these challenges, non-contact techniques, especially optical measurement systems, have been developed. Among these, laser Doppler interferometry is recognised as one of the most reliable and widely used methods. The Laser Doppler vibrometer (*LDV*) represents a state-of-the-art technique for non-contact vibration measurement, offering exceptional sensitivity and resolution for both displacement and velocity [4]. Typically implemented in both laboratory and field environments, *LDVs* exploit the Doppler effect and the principle of optical interferometry to achieve measurement resolutions down to the femtometre scale and bandwidths extending beyond 6 GHz. These performance characteristics are independent of the distance between the sensor and the target, making *LDV* particularly suitable for applications ranging from microstructures to large-scale components. Moreover, as a fully optical and therefore non-invasive technique, *LDV* is ideal for probing delicate or lightweight structures without introducing mechanical loading effects. As anticipated, the operating principle is based on the Doppler effect, whereby a laser beam reflected from a moving surface experiences a frequency shift f_S, related to the surface velocity v and the laser wavelength λ according to the expression $f_S = 2v/\lambda$. To retrieve the velocity information from this frequency shift, *LDVs* employ a laser interferometer.

The basic optical configuration, schematically illustrated in Figure 6, includes a laser beam that is split into two paths: a reference beam and a measurement beam. The measurement beam is focused onto the vibrating surface and, upon reflection, is recombined with the reference beam at the detector. Because the optical path of the reference beam r_2 remains constant over time, any change in the path of the measurement beam $r_1(t)$ due to surface motion gives rise to a modulated interference signal at the detector. This interference pattern is governed by the relation:

$$I_{tot} = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\left(\frac{2\pi(r_1 - r_2)}{\lambda}\right). \tag{1.4}$$

Where I_1 and I_2 are the intensities of the individual beams. Each complete light–dark cycle corresponds to a displacement of half the laser wavelength. To resolve the direction of motion, since movement toward or away from the sensor produces identical modulation patterns, an acousto-optic modulator (Bragg cell) is introduced in the reference path, shifting the laser frequency by a fixed value. This generates a baseline modulation at the detector even when the sample is stationary. Motion towards or away from the detector increases or decreases the modulation frequency respectively, thus enabling unambiguous measurement of both velocity magnitude and direction. Depending on the application, *LDVs* can be configured to demodulate either displacement or velocity. Displacement demodulation, achieved by counting interference fringes, is particularly effective at low frequencies. In contrast, velocity demodulation, derived directly from the Doppler frequency, is more suitable for higher-frequency vibrations. In such cases, the peak velocity of a harmonic signal is given by $v = 2\pi f s$, where f is the vibration frequency and s is the displacement amplitude.

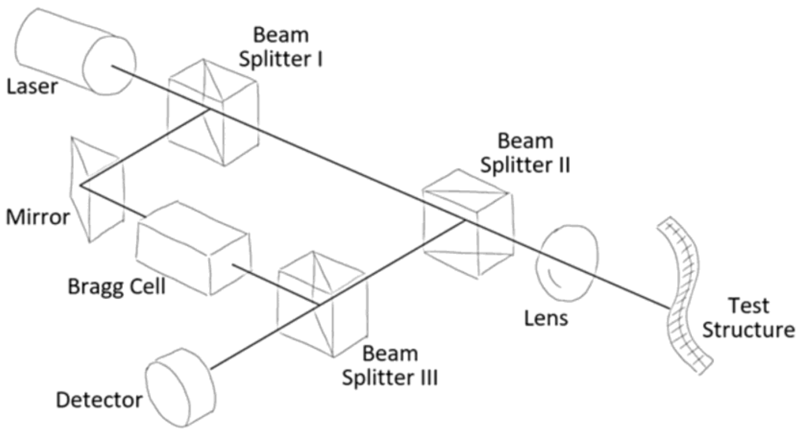


Figure 6: sketch of basic optical configuration of a laser doppler vibrometer

1.3 CLASSICAL ACOUSTIC MEASUREMENT TECHNIQUES

Sound recording and measurement have a long history, dating back to 1857 when Edouard-Léon Scott de Martinville patented the "*phonautograph*". The first electrical microphone was developed by David Hughes for use in telephony, marking a significant advance in acoustic transduction. The condenser microphone, invented by Christopher Wente in 1916, established a design principle that remains fundamental to the majority of scientific measurement microphones in use today. This subsection provides an overview of the transduction principles and performance characteristics of typical acoustic measurement sensors, such as microphones, acoustic cameras, intensity probes and others. Experimental research has demonstrated that the smallest root-mean-square acoustic pressure perceptible on average by young, healthy listeners at $1\ kHz$ is approximately $20\ \mu Pa$. This threshold has been internationally adopted as the reference level for acoustic measurements in air. Acoustic measurements can be performed either *in situ*, under real operational conditions, or within specially designed rooms that provide controlled environments [9,10], such as reverberant or anechoic chambers. A reverberant room is constructed with non-parallel walls and surfaces designed to maximize sound diffusivity within the space, thereby approximating a uniform sound field ideal for measurements of sound power or absorption coefficients. In contrast, an anechoic room is engineered to suppress sound reflections and echoes by using highly absorptive linings on all surfaces. It is often also structurally isolated to block external noise and vibrations. This combination ensures that only direct sound waves are present at the measurement location, effectively simulating free-field conditions and enabling precise characterization of a source's acoustic emission without interference from reflected energy. Both types of facilities are highly specialised and can be particularly expensive to design, construct, and maintain**,** given the stringent requirements for their acoustic performance and isolation. In recent years, there has been growing research interest in acoustics, leading to the development of new types of acoustic facilities, including small anechoic rooms designed for more flexible or cost-effective testing. These smaller chambers can be equipped with active noise control systems that effectively cancel low-frequency sound reflections from the walls [11–14].

CONDENSER MICROPHONE

Most measurement microphones operate on the condenser microphone principle. As shown in Figure 7, in this design, the microphone capsule features an extremely thin diaphragm on its outer surface, which is typically protected by a grid cap that can also be engineered to shape the microphone directional response. The diaphragm and a rigid backplate together form a capacitor. In traditional designs, this capacitor is polarised by an externally supplied direct current (*DC*) voltage.

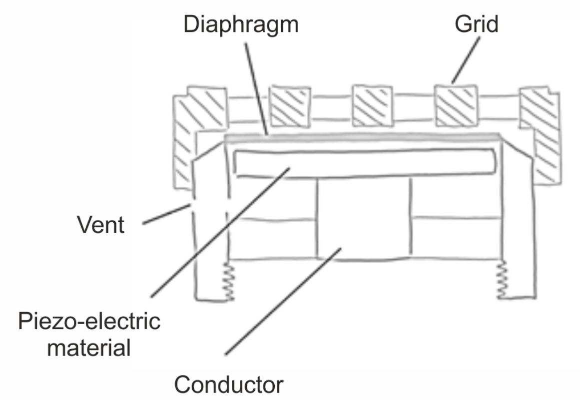


Figure 7: sketch of a condenser microphone

The backplate usually includes a pattern of holes specifically intended to dampen the mechanical resonance of the diaphragm. When exposed to a sound field, the acoustic pressure causes the diaphragm to move relative to the backplate, thereby altering the capacitance of the system. Since the electric charge on the capacitor remains constant, this change in capacitance translates into a corresponding variation in voltage, which can then be measured and analysed. These microphones generally exhibit a uniform directional response over a broad frequency range, and are typically classified as omnidirectional measurement microphones. There are, however, several applications in which a conventional condenser microphone is unsuitable. For example, standard microphone capsules are too delicate to operate reliably in high-temperature environments. A practical solution is offered by probe microphones, which feature a sampling end formed by a small-diameter tube. In this design, the actual microphone element is positioned at a convenient distance from the probe tip, allowing measurements to be taken in challenging conditions while protecting the sensitive capsule.

MICROPHONE ARRAYS OR ACOUSTIC CAMERAS

In many practical scenarios, it is essential not only to measure sound but also to locate, identify, characterise, and quantify the sources responsible for noise generation. A spatial array of microphones addresses this need by simultaneously sampling the radiated sound field at multiple spatial locations. Microphone array methods generally fall into two broad categories, each defined by its underlying source-model assumptions and signal-processing approaches: beamforming and inverse techniques [15,16]. The most basic implementation of beamforming involves a linear array of uniformly spaced microphones. This configuration typically assumes either a single, spatially compact source or multiple, well-separated sources that are mutually uncorrelated, and located at distances much greater than the length of the array. Linear beamformers are especially effective for localising outdoor noise sources over significant distances. When a plane wavefront arrives at the array from a given angle, the difference in arrival times at adjacent microphone pairs depends on the speed of sound, the angle of incidence, and the inter-microphone spacing. However, this simple line-array approach suffers from angular ambiguity and limited directional resolution. For acoustic wavelengths smaller than twice the microphone spacing, spatial aliasing occurs, introducing ambiguity. A common solution to this problem involves sparse arrays with varied microphone spacings, enabling coverage of a broader frequency range without aliasing [17,18]. Beyond linear arrays, many $2D$ array geometries have been developed to improve spatial resolution and flexibility. Examples include *+ shapes* [19], *X shapes*, *star shapes* [20], and *spiral shapes* [21]. Beamforming methods [17,18,22,23] are highly effective for scanning an area to detect discrete, compact sources. Nevertheless, they have limited capacity to resolve the detailed spatial distribution and statistical properties of extended or distributed sources.

As an alternative, inverse techniques [19,24,25] have been developed to characterise and quantify sound fields by estimating the spatial distribution of source strength. These methods offer the advantage of mapping the strength of spatially extended radiating sources but require sophisticated signal-processing algorithms to deliver reliable results. Inverse techniques are particularly suited to cases where the locations of individual sources are known, but their specific strengths are not. Typically, an array of microphones is positioned at some distance from the source region. In principle, the number of microphones should match the number of sources. For stationary (time-invariant) source systems, it is even possible to reduce hardware costs by designating one master microphone position and moving a second microphone sequentially to build up a virtual array. Source strengths are then estimated by inverting the transfer-function matrix and multiplying it by the measured microphone-signal vector, hence the term inverse technique.

SOUND INTENSITY PROBES

Pressure is a scalar quantity without an associated direction, and single-sensor measurement microphones generally exhibit an omnidirectional or uniform directional sensitivity across a broad frequency range. However, in many practical situations, it is necessary to identify dominant noise sources, locate them spatially, and characterise their strength, temporal and spectral properties, as well as the underlying physical mechanisms responsible for their generation.

Sound waves transport energy in two complementary forms: the kinetic energy associated with particle velocity and the potential energy linked to variations in density or pressure. The instantaneous flow of acoustic energy through a unit area is described by the sound intensity [26]. Unlike pressure, sound intensity $\boldsymbol{I}$ is a vector quantity defined as the instantaneous product of the acoustic pressure p and the particle velocity vector $\boldsymbol{u}$ in the medium, and it points in the direction of the particle velocity:

$$\boldsymbol{I} = p\boldsymbol{u}. \tag{1.5}$$

While measuring the instantaneous pressure at a point in a sound field is relatively straightforward using a standard pressure microphone, the major challenge in intensity measurement lies in simultaneously determining the instantaneous particle velocity at the same spatial location. To address this challenge, intensity probes were developed [26,27]. These instruments can be broadly classified into two main types: $p - p$ probes (pressure–pressure) and $p - u$ probes (pressure–particle velocity). For instance, Fahy was the first to propose the use of two microphones and to calculate the sound intensity using the cross-spectral density method [28].

The first reliable commercial system for measuring sound intensity was introduced in 1980. It employed a $p - p$ probe design based on a linearised form of Euler's equation, which relates the particle velocity at a point within a sound field to the spatial gradient of the acoustic pressure:

$$\boldsymbol{u} = -\frac{1}{\rho_0}\int (\boldsymbol{\nabla} p)dt. \tag{1.6}$$

Where ρ_0 represents the mean density of the medium and $\boldsymbol{\nabla}$ denotes the spatial gradient operator. An approximation of this pressure gradient can be obtained by measuring the difference in pressure between two closely spaced points and dividing by their separation distance. In free-field conditions, sufficiently far from the source, the sound intensity is approximately proportional to the mean-square sound pressure. Therefore, if two microphones are positioned closely together at locations 1 and 2, separated by a distance Δr, the particle velocity can be approximated by:

$$u = -\frac{1}{\Delta r \rho_0}\int_{-\infty}^{t} \big(p_1(\tau) - p_2(\tau)\big)d\tau. \tag{1.7}$$

The separation Δr must be small relative to the acoustic wavelength to ensure the validity of this approximation, but not so small that the pressure difference falls below the threshold of accurate measurement. The acoustic pressure at the evaluation point is typically estimated as the average of the two measured pressures. A conventional two-microphone intensity probe, referred to as a $p-p$ probe, consists of two nominally identical pressure microphones mounted face to face, separated by a rigid coaxial spacer (as illustrated in Figure 8). The operational frequency range of such a system is constrained not only by the microphones sensitivity and the electronic noise floor of the microphones and preamplifiers, but more fundamentally by the spacing between the microphone diaphragms. At high frequencies, the acoustic wavelength becomes too short for the pressure-difference approximation to remain valid. Conversely, at low frequencies, the pressure difference between the two microphones becomes too small, causing the derived particle velocity signal to be increasingly susceptible to system noise.

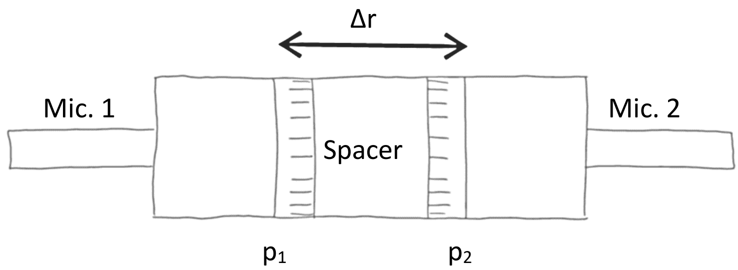


Figure 8: sketch of a conventional p-p intensity probe

The second principal type of intensity probe is the so-called $p-u$ probe, which uses a direct approach to measure acoustic particle velocity. A widely adopted design incorporates two extremely thin resistance wires positioned close together and parallel to each other, which are electrically heated to temperatures exceeding 200 °C. These wires are suspended across a gap in the probe structure, allowing the incident sound field to pass transversely through them.

The wires are mounted between two parallel cylindrical bodies that serve to amplify the particle velocity of the sound wave in the measurement region. When heated, the temperature distribution across the plane of the wires results from the combined fields around each wire. Crucially, the downstream wire develops a higher temperature peak than the upstream wire due to convective effects in the moving air. This temperature difference leads to a differential change in the electrical resistance of the wires, which is directly proportional to the local particle velocity across them. The direction of the velocity vector is indicated by the sign of this resistance difference.

For intensity measurements, the $p - u$ probe design also includes a small pressure microphone co-located within the same measurement channel as the heated wires. By multiplying the signals from the pressure sensor and the particle velocity measurement, the probe yields the component of the sound intensity vector aligned with the measurement axis.

Because particle velocity is inherently directional (a vector quantity), the $p - u$ configuration offers strong advantages in suppressing spurious noise and unwanted reflections. This makes it particularly effective in array-based applications for the localisation and quantification of sound sources.

NEARFIELD ACOUSTIC HOLOGRAFY (*NAH*)

Near-field acoustic holography (*NAH*) is a measurement technique that leverages the wave nature of sound to reconstruct and visualise three-dimensional acoustic fields from measurements taken on a two-dimensional surface close to the source. The method is conceptually rooted in optical holography, which relies on interference and diffraction phenomena to record a complete wavefront on a flat surface using a coherent light source. In acoustics, a similar principle is applied: by measuring the sound pressure over a plane in the near field, it becomes possible, through numerical back-propagation based on known Green’s functions and the wave equation, to recover not only the three-dimensional sound pressure field but also derived quantities such as particle velocity, acoustic intensity vectors, and the surface velocity or intensity distribution of vibrating sources.

The fundamental strength of *NAH* lies in its high information content: measurements on a limited, planar surface (the so-called hologram plane) can be used to characterise the entire three-dimensional wave field generated by scattering bodies or active sources. Traditionally, *NAH* methods have been applied to stationary sound fields at single frequencies. However, many practical engineering sources exhibit non-stationary, time-varying behaviour, which requires the development of advanced *NAH* techniques capable of capturing and analysing these time-dependent properties [29–35]. Despite its relative conceptual simplicity, near-field acoustic holography is a powerful measurement approach because it fully exploits the governing wave equation to achieve detailed spatial reconstruction of acoustic phenomena.

1.4 OPTICAL CAMERA MEASUREMENTS

Camera measurements are non-contact optical techniques that have rapidly advanced in recent decades, driven by improvements in computing power, memory, and high-resolution camera sensors. Generally, optical measurement techniques can be broadly categorized into laser-based approaches and passive-light, image-based systems. Among the latter, photogrammetry is a prominent technique for extracting geometry, displacement, and deformation data from photographs or digital images, building on principles of perspective established as early as Leonardo da Vinci [36]. Initially applied to military reconnaissance, photogrammetry has expanded with digital cameras into civil and mechanical engineering, taking forms such as close-range photogrammetry, dynamic photogrammetry, machine vision, and robot vision [37]. Two-dimensional ($2D$) photogrammetry relies on a single fixed camera and is limited to in-plane measurements, while to measure $3D$ vibrations at least two camera views, observing matching points of interest, are required. This is often achieved using a synchronized pair of high-speed cameras and to avoid possible synchronization and optical errors in multi-camera setups, the so-called single-camera multi-view methods are currently studied actively. These methods often utilize additional optical elements such as mirror systems to project multiple views of the structure surface onto a single camera sensor, reducing the measurement resolution as a result [38]. The frequency-domain triangulation method [39] is another technique that enables the use of a single camera for spatial vibration measurements, which, exploits full image-sensor resolution. It can be implemented with a still-frame camera and does not require synchronization of multiple cameras [40]. Therefore, stereovision and multi-camera systems enable three-dimensional measurement of complex or curved surfaces [36,41]. The measurement of structural stress and strain using photogrammetry was first introduced by Peters and Ranson in the 1980s [42], and further developed through Digital Image Correlation (*DIC*) methods by Sutton, Chu, and colleagues [43,44], combining imaging with numerical techniques to quantify full-field deformation. These techniques use speckled patterns to track surface changes across time, with broad applications in solid mechanics [45], aerospace [46], $3D$ shape measurement [47], and infrastructure inspection [48]. In the last decades, more computationally efficient techniques, such as the simplified optical flow method [49,50] have also been developed to be used when fast postprocessing is required. High-speed camera improvements have also made it possible to use these approaches for vibration measurements, enabling non-contact, full-field dynamic characterization of structures. A key part of camera-based measurement is the point-extracting process, which can use two main methods. The artificial marker method (or target-based approach) involves placing high-contrast optical targets such as dots or speckles on the object, with cameras recording their positions for precise analysis via image processing algorithms [42]. By contrast, the natural feature method (untargeted approach) uses surface features of the object as tracking regions, eliminating the need for markers and improving efficiency by reducing effort in applying targets on the test surface. Both approaches rely on advanced optical algorithms, mainly point tracking and digital image correlation. Point tracking focuses on identifying and following specific feature points or markers over time, allowing precise motion trajectory analysis without relying on grayscale patterns. *DIC*, on the other hand, analyses grayscale variations in defined image regions to compute surface deformation and displacement fields, enabling full-field measurement and detailed deformation mapping [51]. These optical techniques, leveraging advanced image processing and computer vision methods, play a critical role in engineering applications that demand high-precision measurements, such as structural health monitoring and vibration analysis. In this doctoral thesis, two types of cameras were used: a *high-speed camera*, which is a conventional frame-based imaging system, and *event-based cameras*, which record only changes in light intensity rather than capturing full frames at fixed intervals. The specific characteristics of these cameras, along with the post-processing methods employed to extract vibration information from the recorded data, will be presented in detail in dedicated chapters of this thesis.

1.5 MOTIVATION OBJECTIVES AND CONTRIBUTION OF THE THESIS

In recent years, research in the field of acoustic measurements has focused increasingly on developing simpler and more practical approaches for characterizing the full field vibration and sound radiation from machinery and structural elements. Traditionally, such measurements are conducted in dedicated facilities such as large reverberant or anechoic chambers [9,10,52,53]. These facilities, however, are expensive to build and operate, and are typically available only at certification or research centres that provide measurement services to third parties. When it is difficult or too costly to move the machinery or structure under investigation, measurements are therefore often performed *in situ*, directly at the installation site [9,10,54]. This approach, however, is delicate to implement because it requires careful arrangements to minimize the effects of sound reflections from walls or partitions, as well as flanking noise from other machines or systems [55–58], for example, Rohlfing et al. [59] demonstrated the reconstruction of the sound radiation produced by the flexural vibration of a flat panel using pointwise measurements acquired with a scanning laser vibrometer. Their method involved discretizing the surface into a grid of elemental radiators [58,60] and measuring transverse velocities at the centres of these elements to compute the radiated sound field using a discretized version of the Kirchhoff–Helmholtz integral equation. This concept has been further developed to represent sound radiation in terms of radiation modes, which provide deeper insight into the physical mechanisms of sound emission [61,62]. Building on this idea, the present doctoral research proposes an alternative, camera-based approach to reconstruct the sound radiation field produced by the flexural vibrations of structures. This method resembles the elemental radiator technique but uses high-speed cameras to capture the vibration field [63] instead of relying on a scanning laser vibrometer [59,61,62]. Like the *LDV* method, it is largely unaffected by the acoustics of the measurement environment, flanking sources, or background noise. However, unlike the *LDV*, the acquisition time is significantly shorter, which makes full-field, *in situ* measurements more practical. While these factors can still indirectly influence structural vibrations, this effect is expected to be negligible in air, in large spaces with low reverberation, and under typical practical conditions. Camera-based measurements therefore enable fast, dense sampling of the vibration field, allowing for higher accuracy in predicting the radiated sound field without the need for an anechoic chamber. Previous experimental work [64] has also shown that, in contrast to scanning vibrometer, only short video recordings are required to reconstruct the entire vibration field, thereby enabling measurements to be carried out under nominal operating conditions and avoiding possible alterations in the structural dynamics, such as those induced by temperature variations during prolonged acquisition times. Finally, with the goal of reducing the amount of data stored during acquisition, enabling the use of more compact and convenient sensors suitable for installation even in hard-to-reach locations, while still maintaining high dynamic range and effective temporal resolution, this thesis also explores the use of novel sensors known as neuromorphic sensors or event cameras [65–69]. These sensors were investigated specifically for measuring the surface vibrations of structures, laying the groundwork for future extensions of this methodology toward sound radiation estimation.

In summary, the principal objectives of the study are:

- To develop an optical *camera-based* approach for the full-field reconstruction of the flexural vibration and sound radiation of distributed structures including closed shells;
- To introduce a new measurement approach of full-field vibrations of flexible structures based on a new type of camera that generates fast, "*light-weight*" streams of image information called events.

Also, the main contributions of this doctoral thesis are:

- The derivation of the full-field flexural response of vibrating structures from high-speed camera measurements, and frequency domain triangulation image post-processing;
- The derivation of the sound radiation field of a baffled thin-shell cylinder from high-speed camera measurements, including the calculation of acoustic maps

and the determination of sound power levels in narrow-band spectra and one-third-octave bands. In this context, a novel strategy for acoustic power calculation based on collapsing the acoustic measurement surface onto the radiating surface is introduced and applied to this analytical case;

- The extension of the same collapsed-surface strategy to a more complex geometry, a prismatic thin-shell structure is also produced. Here, a Boundary Element Method (*BEM*) model is employed to derive the acoustic radiation from high-speed camera measurements, and the results obtained with the standard approach are compared with those from the novel collapsed-surface method, for both narrow-band and one-third-octave representations;
- The extraction of deflection shapes of a cantilever beam from optical measurements carried out with event cameras.

1.6 STRUCTURE OF THE THESIS

This thesis is divided into nine chapters. The first four chapters provide a comprehensive overview of the theoretical formulations, experimental principles, and *state-of-the-art* tools that form the foundation for the research conducted in this work. To start with, Chapter 1 briefly introduces the physics of mechanical vibrations and acoustics, presenting classical measurement methods and briefly touching upon optical measurement techniques. Chapter 2 then describes the fundamental design principles of conventional frame-based cameras and event-based cameras. It also highlights the main post-processing techniques used for image analysis with both types of sensors. Chapters 3 and 4 provide the methodological foundations of the work. The first introduces the computer vision techniques employed for vibration measurement, including marker tracking and triangulation methods, while the second recalls the essential principles of acoustics relevant to this work, from the wave equation to the Boundary Element Method, which are later expanded and used to derive the radiated sound power. Chapters 5 to 9 then present the original formulations and experimental developments proposed in this thesis, where the methods introduced are derived, implemented, and validated through different test cases. To start with, Chapter 5 applies the tools presented in the early chapters to the measurement of vibrations using a single high-speed camera. After describing the test structures and setups, the frequency-domain triangulation method is used to identify structural modes and to derive vibration fields and kinetic energy spectra. Building on this, Chapter 6 focuses on the acoustic radiation of a baffled cylinder, where microphone array measurements and the Kirchhoff–Helmholtz formulation allow the calculation of acoustic maps and sound power levels. The analysis is then extended in Chapter 7 to a small prismatic box, enabling a comparison between different structural geometries through the derivation of sound pressure fields and radiated power. Chapter 8 moves to event-based sensing, showing how synchronous event cameras, combined with neural network frame reconstruction and triangulation, can be used to extract the deflection shapes of a cantilever beam. Finally, Chapter 9 gathers the conclusions of this work and highlights perspectives for future developments.

2

FRAME-BASED VS. EVENT CAMERAS

Computer vision techniques aim to extract quantitative information from images, such as spatial coordinates and displacements [70,71]. In this context, camera models are essential to establish the relationship between $3D$ points on an object surface and their $2D$ projections in the image plane. This chapter introduces the fundamental pinhole camera model, which underlies both conventional frame-based imaging systems and event-based sensors. The chapter then describes the main characteristics of high-speed cameras and event cameras, highlighting their differences in terms of acquisition principles.

2.1 PINHOLE MODEL

The most basic geometric model of image formation is the pinhole camera, schematically illustrated in Figure 9. In this model, a small aperture (the pinhole) allows a single light ray from each point in the scene to pass through, forming an image on the opposite surface known as the image plane.

Consider a point M in the $3D$ scene with coordinates (X, Y, Z) and let M' with coordinates (X', Y', Z') denote its projection onto the image plane through the pinhole C. If f represents the distance from the pinhole to the image plane (the focal length), similar triangles yield the relationship:

$$\begin{cases} X' = -\dfrac{fX}{Z} \\ Y' = -\dfrac{fY}{Z} \\ Z' = -f \end{cases} \tag{2.1}$$

These equations describe perspective projection, characterized by an inversion of the image with respect to the scene (both left–right and top–bottom) as indicated by the negative sign. The geometric model of the camera adopted here is often called the pinhole or perspective model, consisting of the image plane Q and the center of projection (COP) C, located at a distance f (focal length) from the image plane. The collection of lines passing through C forms the projective bundle. The line through C orthogonal to Q is the principal axis of the camera, with its intersection on Q termed the principal point. The plane F, parallel to Q and containing C, is called the focal plane, while points lying in F project to infinity on Q. To analytically describe the perspective projection performed by the camera, appropriate Cartesian coordinate systems are introduced to represent points in three-dimensional space and their projections on the image plane. In a simplified scenario, the world (or object) coordinate system is chosen to coincide with the camera reference frame, defined as a right-handed triplet (X, Y, Z) centered at C, with the $Z - axis$ aligned along the principal axis. Additionally, a reference frame (u, v) is defined on the image plane Q, centered at the principal point, with axes u and v aligned with X and Y respectively. From the similarity of triangles, the following relation is obtained:

$$\begin{cases} u = -\dfrac{f}{Z}X \\ v = -\dfrac{f}{Z}Y \end{cases} \tag{2.2}$$

This transformation in Cartesian coordinates is nonlinear due to the division by Z. In homogeneous coordinates, the projection becomes linear. Defining: $\boldsymbol{m} = [u, v, 1]^T$ and $\boldsymbol{M} = [X, Y, Z, 1]^T$ as the homogeneous coordinates of the image point and the 3D point respectively (note that setting the last coordinate to 1 excludes points at infinity; to include them, a generic third component should be used), the perspective projection equation can be written as:

$$\boldsymbol{m} \simeq \boldsymbol{PM} \tag{2.3}$$

Where $\simeq$ denotes equality up to scale, and $\boldsymbol{P}$ is the perspective projection matrix (*PPM*) or camera matrix describing the geometric model of the camera.

The focal length f determines the magnification of the image: placing the image plane closer to the pinhole reduces the image size, while increasing f enlarges it. However, moving the image plane further away also narrows the field of view, reducing the solid angle encompassing the visible portion of the scene. The division by Z in the projection formula introduces the foreshortening effect**,** in which objects appear smaller in the image as their distance from the camera increases. For instance, railway sleepers, though equal in length in reality, appear progressively shorter in the image depending on their distance from the observer. When the object being imaged is relatively thin compared to its average distance from the camera, the perspective projection can be approximated by an orthographic projection. Specifically, if the depth Z of object points varies within a range $Z_0 \pm \Delta Z$ satisfying $\Delta Z / Z_0 \ll 1$, the scale factor f/Z can be approximated by the constant f/Z_0. The resulting projection equations simplify to:

$$\begin{cases} X' = -\dfrac{f}{Z_0} X \\ \\ Y' = -\dfrac{f}{Z_0} Y \end{cases} \tag{2.4}$$

representing an orthographic projection scaled by f/Z_0. This approximation was famously recommended by Leonardo da Vinci whenever $\Delta Z / Z_0 < 1/10$.

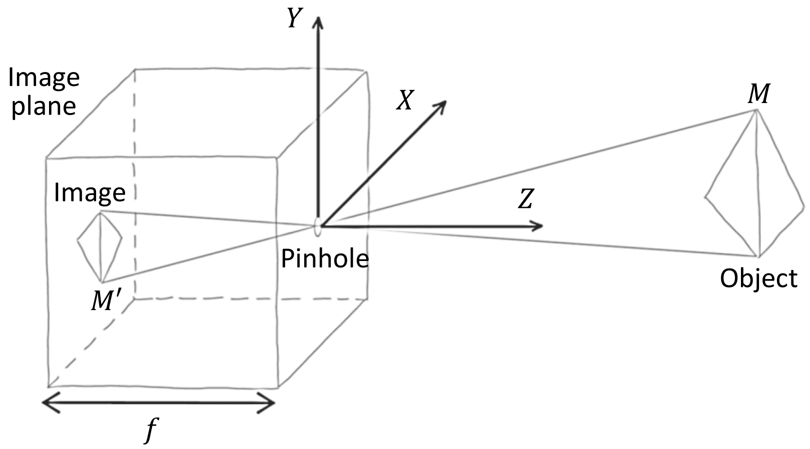


Figure 9: pinhole camera model

2.2 INTRINSIC PARAMETERS

The transformation from sensor coordinates to pixel coordinates is modelled by an affine transformation that accounts for the translation of the origin to the upper-left corner, independent scaling of the u and v axes to convert units from meters to pixels, and the inversion of the axes directions. This transformation is represented by the matrix:

$$\boldsymbol{A} = \begin{bmatrix} -k_u & 0 & u_0 \\ 0 & -k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.5}$$

where (u_0, v_0) are the pixel coordinates of the principal point, and k_u, k_v are the inverses of the effective pixel sizes along the u and v directions, respectively. Matrix $\boldsymbol{A}$ converts coordinates from meters to pixels and premultiplies the perspective projection matrix (*PPM*), resulting in:

$$P = AP_{persp} = K[I_3 \quad \mathbf{0}], \tag{2.6}$$

where:

$$K = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.7}$$

is the intrinsic parameter matrix, and $[I_3 \quad \mathbf{0}]$ encodes the essential perspective transformation without parameters. This intrinsic matrix $\boldsymbol{K}$ relates normalized image coordinates, expressed in metric units relative to the camera, to pixel coordinates. The inverse relation is given by:

$$\boldsymbol{q} = \boldsymbol{K}^{-1}\boldsymbol{m}, \tag{2.8}$$

where $\boldsymbol{m}$ denotes pixel coordinates and $\boldsymbol{q}$ normalized coordinates.

The most general form of $\boldsymbol{K}$ also includes a skew angle θ, which represents the angle between the u and v axes, typically $\pi/2$. Defining the focal length in horizontal pixels as $\alpha_u = fk_u$ and the pixel aspect ratio as $\tau = k_v/k_u$, the intrinsic matrix can be expressed as:

$$K = \begin{bmatrix} \alpha_u & -\alpha_u/tan\theta & u_0 \\ 0 & \tau\alpha_u/sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.9}$$

This matrix has five degrees of freedom, corresponding to the intrinsic parameters α_u, u_0, v_0, τ and θ. Common simplifications assume orthogonal axes and square pixels, reducing the number of parameters. Furthermore, although physically the image plane is positioned behind the pinhole (yielding a negative focal length $-f$ and an inverted image) the digital image formation process generally adopts a positive focal length $+f$, producing non-inverted images. Consequently, in the digital model, the diagonal elements K_{11} and K_{22} are negative.

2.3 ESTRINSIC PARAMETERS

In general, the world reference frame does not coincide with the camera reference frame. To model this, a change of coordinates is introduced, consisting of a rigid transformation: a rotation $\boldsymbol{R}$ followed by a translation $\boldsymbol{t}$. Let $\boldsymbol{M_C}$ denote the homogeneous coordinates of a $3D$ point in the camera coordinate system, and $\boldsymbol{M}$ its homogeneous coordinates in the world coordinate system. Their relationship is given by:

$$\boldsymbol{M_C} = \boldsymbol{GM}, \tag{2.10}$$

where:

$$G = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix}. \tag{2.11}$$

The matrix $\boldsymbol{G}$ encodes the extrinsic parameters of the camera, describing its position and orientation relative to the world. Specifically, $\boldsymbol{R}$ represents the rotation, while $\boldsymbol{t}$ represents its position. By parameterizing $\boldsymbol{R}$ using the Euler angles, the camera external orientation is fully described by six extrinsic parameters: three rotation parameters and three translation components (t_1, t_2, t_3).

The perspective projection matrix combines both intrinsic and extrinsic parameters and from Equation (2.6) and (2.10) it is possible to obtain:

$$\boldsymbol{m} \sim \boldsymbol{K}[I_3 \quad \boldsymbol{0}]\boldsymbol{G}\boldsymbol{M} \tag{2.12}$$

so that the overall PPM becomes:

$$\boldsymbol{P} = \boldsymbol{K}[\boldsymbol{R} \quad | \quad \boldsymbol{t}]. \tag{2.13}$$

Here, $\boldsymbol{K}$ contains the intrinsic parameters (describing the camera internal geometry), while $[\boldsymbol{R} \quad | \quad \boldsymbol{t}]$ models the extrinsic parameters, describing the camera orientation in the world. Importantly, the *PPM* is only defined up to scale: multiplying $\boldsymbol{P}$ by any nonzero scalar λ yields the same projection. As a result, although $\boldsymbol{P}$ has 12 entries, it has only 11 degrees of freedom in total (five intrinsic parameters and six extrinsic parameters).

2.4 RADIANCE

Light emitted by sources such as the sun interacts with the surfaces of objects in the scene: part of it is absorbed, while part is scattered in various directions. Among these, only the rays directed toward the camera contribute to the image formation process.

The brightness $I(\boldsymbol{m})$ of a pixel $\boldsymbol{m}$ in the image is proportional to the amount of light scattered toward the camera from a surface element S_M centred at point $\boldsymbol{M}$, which projects onto the pixel. This scattered light depends both on the reflectance properties of the surface and on the distribution and orientation of the illumination. Formally, the quantity of light energy propagating along a ray is described by the concept of radiance. Radiance $L(\boldsymbol{M}, \boldsymbol{v})$ is defined as the power of light per unit surface area per unit solid angle, emitted or received at point $\boldsymbol{M}$ in direction $\boldsymbol{v}$. If we neglect atmospheric attenuation, the outgoing radiance L_0 from point $\boldsymbol{M}$ in the direction of the camera coincides with the incoming radiance received by the image plane at the corresponding location $\boldsymbol{m}$:

$$I(\boldsymbol{m}) = L_0(\boldsymbol{M}, \boldsymbol{v}). \tag{2.14}$$

The material reflective properties are described by its Bidirectional Reflectance Distribution Function (*BRDF*) [72], denoted as $\rho_{BRDF}(\boldsymbol{M}, \boldsymbol{v}, \boldsymbol{s})$. This function expresses how light arriving from direction $\boldsymbol{s}$ is reflected in direction v, and depends on the surface geometry at point M. In general, the *BRDF* is a four-dimensional function parameterized by the angles of $\boldsymbol{v}$ and $\boldsymbol{s}$, although for many materials it can be simplified to a function of the angles formed with the surface normal $\boldsymbol{n}$, thus reducing the dimensionality.

The reflectance equation expresses the outgoing radiance $L_0(\boldsymbol{M}, \boldsymbol{v})$ as an integral over the hemisphere Ω above the surface, accounting for all incoming light directions:

$$L_0(\boldsymbol{M}, \boldsymbol{v}) = \int_{s \in \Omega} \rho_{BRDF}(\boldsymbol{M}, \boldsymbol{v}, \boldsymbol{s}) L_i(\boldsymbol{M}, \boldsymbol{s})(\boldsymbol{n}^T \boldsymbol{s}) ds. \tag{2.15}$$

$(\boldsymbol{n}^T \boldsymbol{s})$ accounts for the orientation of the surface relative to the incoming direction $\boldsymbol{s}$.

In practice, solving this integral fully would require considering all direct and indirect lighting contributions from the entire environment. However, as a first approximation, it is common to consider only primary sources. If a single distant light source is assumed, the integral simplifies to a product:

$$L_0(\boldsymbol{M}, \boldsymbol{v}) = \rho_{BRDF}(\boldsymbol{M}, \boldsymbol{v}, \boldsymbol{s}) L_i(\boldsymbol{M}, \boldsymbol{s})(\boldsymbol{n}^T \boldsymbol{s})[\![\boldsymbol{n}^T \boldsymbol{s} > 0]\!]. \tag{2.16}$$

Among the possible *BRDF* models, the simplest is the Lambertian or diffuse model, which assumes that light is scattered equally in all directions. In this case, the radiance $L_0(M)$ does not depend on the viewing direction, and the surface appears equally bright from any angle. The Lambertian *BRDF* is constant and determined by the albedo $\rho_{BRDF}(\boldsymbol{M})$, a value between 0 (completely absorbing) and 1 (perfectly reflective).

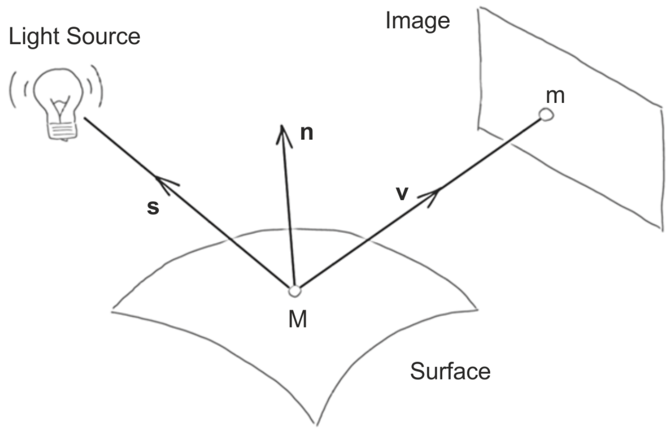


Figure 10: elements involved in the reflectance equation

At the opposite end of the spectrum lies specular reflection, characteristic of mirror-like surfaces, where the reflected light is concentrated along a single direction, such that the angle of reflection equals the angle of incidence, and both rays lie in the same plane. Figure 10 illustrates the elements involved in the reflectance equation, showing the surface point $\boldsymbol{M}$, its normal $\boldsymbol{n}$ and other parameters.

2.5 HIGH-SPEED CAMERAS

High-speed cameras are frame-based sensors capable of capturing thousands to millions of frames per second, enabling the detailed analysis of rapid phenomena. The *CMOS* (Complementary Metal Oxide Semiconductor) is a type of active image sensor composed of an array of photodiode–amplifier pairs that together form the individual picture elements. The output signal consists of an amplified voltage on a readout capacitor, generated by the charge produced when the photodiode is exposed to light [73]. In colour image sensors, each image element comprises three separate sensors, one each for red, green, and blue wavelengths, while in monochrome cameras, each image point is formed by a single photosensitive element. These elements are arranged in a regular grid and each element independently detects image information, resulting in a raster image that, in the case of a grayscale image, can be represented as a matrix of measured light-intensity values at individual points.

The number of image elements per row (n_u) and per column (n_v) depends on the ratio between the physical dimensions of the image sensor (S_u, S_v) and the size of each photosensitive element (E_u, E_v):

$$\begin{cases} n_u = \dfrac{S_u}{E_u} \\ n_v = \dfrac{S_v}{E_v} \end{cases} \tag{2.17}$$

The optical magnification defines the ratio between the sensor dimensions and the measurement area that can be acquired using the chosen imaging parameters. The size of each photosensitive element on the sensor, combined with the optical magnification, determines the spatial sampling of the measurement area and thereby defines the measurement resolution.

Images can be formally described as multi-parametric functions of position, denoted as $f(u, v)$ [70]. Therefore, a grayscale image is a two-dimensional function $f(u, v)$, where each image element corresponds to an intensity value. It can be visualized as a $3D$ surface over the (u, v), plane, in which the “height” represents the grayscale level at each point as shown in Figure 11. The axes refer to the position of the image element, with the unit typically referred to as a pixel.

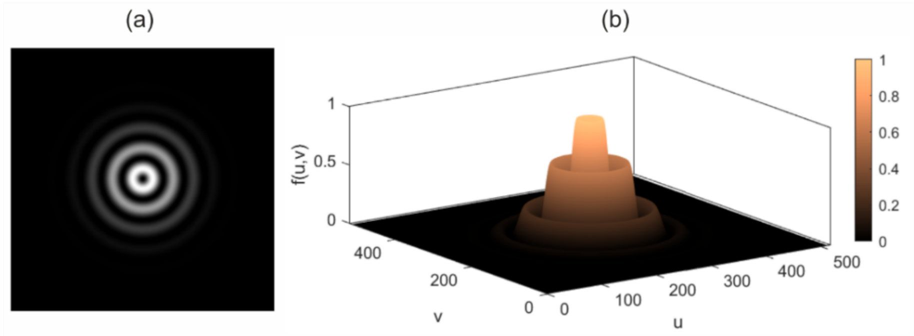


Figure 11: two-dimensional function where each image element corresponds to an intensity value. a) grey scale image, b) 3D surface visualisation of (a)

A two-dimensional digital grayscale image is often represented in discrete form as $f(v_i, u_j)$ where i and j index the pixel rows and columns, respectively. In matrix notation, a digital image of dimensions $n_v \times n_u$ is written as:

$$\begin{bmatrix} f(0,0) & \cdots & f(0, n_u - 1) \\ \vdots & \ddots & \vdots \\ f(n_v, 0) & \cdots & f(n_v - 1, n_u - 1) \end{bmatrix}. \tag{2.18}$$

As light is reflected from the surface of the sample and passes through the optical system to reach the image sensor, its intensity is indirectly measured as an analogue voltage signal. Although this voltage is continuous and well-defined at every instant in time, digital systems operate in discrete domains. Therefore, the analogue signal must be mapped to a digital domain via *Analog-to-Digital Conversion* (*ADC*). The *ADC* process in image sensors consists of two main stages: quantization and sampling. Quantization refers to the discretization of the dependent variable, namely, the grayscale intensity associated with each image element. Sampling, on the other hand, refers to the discretization of the independent variables. In the case of temporal sampling, the independent variable is time t, whereas for spatial sampling in digital images, the independent variables are the image plane coordinates u and v.

QUANTIZATION

The key parameter governing the quantization of the dependent variable is the number of bits used in the encoding process. This number defines how many discrete values the measured quantity can assume. In the context of digital imaging, it specifies the number of discrete grayscale levels that a single image element (pixel) can represent. The bit depth of the sensor sets the maximum number of quantization levels available to describe the intensity of light at each pixel location.

For example, a $16 - bit$ image sensor can encode $2^{16} = 65536$ distinct brightness levels.

TEMPORAL SAMPLING

When measuring high-frequency dynamic processes, the sampling interval Δt is inversely related to the sampling frequency f_s, and represents one of the most critical parameters in configuring the measurement system:

$$f_s = \frac{1}{\Delta t}. \tag{2.19}$$

The continuous signal $f(t)$ is thus represented after sampling by a discrete set of values taken at integer multiples of the sampling interval. In temporal analog-to-digital conversion for imaging data, the sampling rate is determined by the frame acquisition speed of the digital camera. By selecting the sampling interval, one defines the maximum frequency f_{max} that can be accurately represented in the discretized signal, commonly referred to as the Nyquist frequency. This is expressed by the Nyquist criterion $\Delta t \leq 1/2 f_{max}$ [74].

Up to this point, temporal sampling has been described assuming that image acquisition occurs instantaneously. However, in practice, capturing an image requires a finite exposure time long enough to ensure sufficient illumination of the sensor. The amplitude of the signal at time t is not read instantaneously. Instead, the measurement corresponds to the integration of the sensor output voltage over the exposure interval T [73]. An important consequence of this process is that the measured grayscale value for a given image element may be inaccurate, especially in regions of the signal with rapid variations, leading to blurred areas when an inappropriate exposure time is used to image fast-moving objects.

SPATIAL SAMPLING

An analogous concept to the temporal sampling interval Δt, which governs temporal resolution, is the spatial sampling interval Δu and Δv, which determines the spatial resolution of a digital image. The Nyquist criterion applies equally in the spatial domain: accurately capturing patterns with high spatial frequency (high wave number κ) requires sufficiently fine spatial sampling. Under the assumption of isotropy, the minimum sampling interval needed to avoid loss of information is given by:

$$\Delta u = \Delta v = \frac{1}{2\kappa}. \tag{2.20}$$

If the spatial sampling interval is too large aliasing occurs. This manifests in the recorded image as misleading patterns with artificially lower spatial frequency, effectively distorting the true geometry of the object. To minimize aliasing artifacts in experimental imaging, spatial low-pass filtering is commonly employed. Such filtering smooths the image by attenuating high-frequency components, thereby preventing undersampling of fine details. However, this comes at the cost of reduced image sharpness and loss of high-frequency information.

2.6 EVENT CAMERAS

Unlike conventional full-frame sensors, in which radiance is sampled periodically according to an internal clock that governs the acquisition frequency (as described in Section 2.5), the operation of event cameras is fundamentally driven by the observed scene. More specifically, each pixel in an event camera reacts exclusively to variations in radiance over time. This acquisition paradigm is commonly referred to as *scene-driven* or *event-driven*, as it is regulated by asynchronous events generated directly by the sensor underlying circuitry (see Figure 12). According to this principle, a predominantly static scene produces a minimal number of events, whereas a scene rich in rapid motion results in a high volume of events. This amount is roughly proportional to the percentage of pixels detecting a significant change in radiance relative to their previously recorded state.

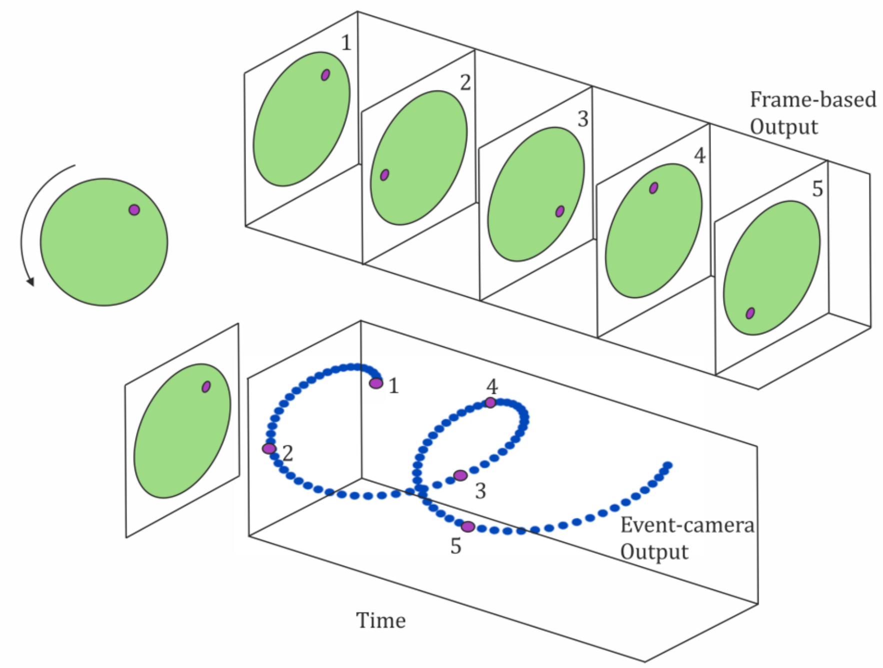


Figure 12: frame-based output vs event-based output of a rotating disk with one marker on the surface

Crucially, each pixel operates independently and generates events solely based on the radiance it detects, without influence from neighbouring pixels, in this way the outcome is an asynchronous stream of events containing only the information relevant to reconstruct the dynamic content of the scene. Consequently, the concept of a *frame* becomes entirely incompatible with this sensing modality. This acquisition strategy leads to a significantly reduced output rate, as pixels that do not perceive sufficient variation remain silent. Importantly, this does not limit the operational frequency of individual pixels; on the contrary, event-based sensors exhibit exceptionally high temporal resolution. Moreover, the reduction in output bandwidth does not compromise the fidelity of the information encoded: by sampling only the variations in radiance, the sensor retains all necessary information to accurately capture and reconstruct motion.

The ability to represent a scene using fewer data while maintaining informational equivalence to that of a traditional camera offers several advantages, particularly in terms of latency and bandwidth. In scenarios requiring transmission over a network, it is desirable to minimize data volume to avoid overloading communication channels (an aspect especially critical in real-time vision applications). Furthermore, power consumption can be a limiting factor in embedded or portable systems, and event cameras are specifically designed to operate with minimal energy consumption.

WORKING PRINCIPLE OF A *DVS* SENSOR

The *Dynamic Vision Sensor* (*DVS*) is stimulated by changes in scene radiance and operates in a fully asynchronous and independent manner at the pixel level. Each pixel continuously monitors the local radiance L and compares its current value with an internal reference:

$$\ln\big(L(x, y, t)\big) - \ln\big(L(x, y, t - \Delta t)\big) = \pm C. \tag{2.21}$$

When the difference between the two exceeds a predefined threshold C (either positively or negatively) the pixel generates an event. Each event encodes several pieces of information: the pixel coordinates (x, y), the timestamp of the event t, and the polarity of the radiance change (positive for ON events, negative for OFF events). A key feature of this sensing mechanism lies in the dynamic update of the comparison threshold: once an event is triggered, the internal reference is reset to the current radiance value, allowing future events to be generated in response to new changes, rather than with respect to a static threshold. This dynamic adaptation ensures responsiveness to temporal variations in radiance, while avoiding redundant event generation. The radiance incident on a pixel generally results from both ambient illumination and reflected light from objects in the scene. Consequently, a variation in ambient radiance might be mistakenly interpreted as object motion. To mitigate this ambiguity, it is typically assumed that ambient illumination remains approximately constant over time. Under this assumption, any radiance variation detected by the sensor is attributed primarily to motion or structural change within the scene [65]. To better illustrate the described behaviour, consider the schematic representation in Figure 13. The vertical axis of the plot shows the logarithmic variation of radiance over time, expressed in volts. Both positive and negative thresholds are marked, along with the associated reset values.

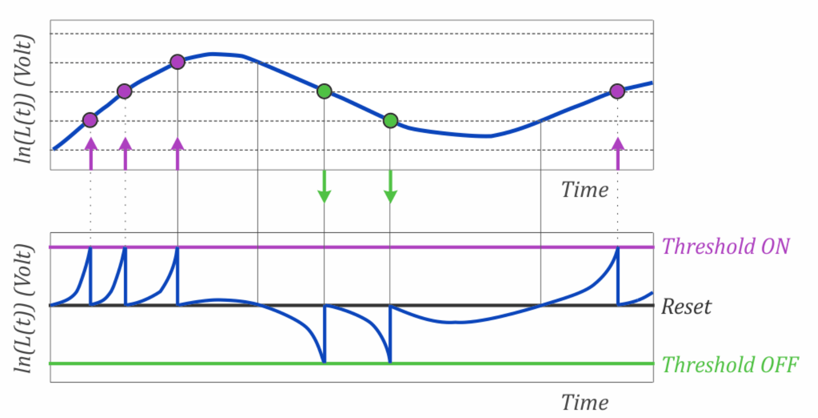


Figure 13: working principle of an event-driven sensor

Suppose the initial threshold level coincides with the first recorded radiance value. After a short interval, the radiance rises to the point where its difference from the reference exceeds the positive threshold, triggering an ON event. Each time this condition is met, a new event is generated, and the threshold is updated accordingly. Eventually, the radiance curve reaches a plateau and then begins to decrease. In this second phase, once the radiance drops sufficiently below the reference, OFF events are generated. As can be observed from the figure, the frequency of the triggered events is not constant, moreover it is important to clarify that the rate of change in radiance does not influence whether an event will be generated, but rather determines the frequency of event generation. Once an event is generated by the sensor, it must be transmitted downstream for further processing. However, most conventional electronic communication interfaces are designed under the assumption of synchronous operation, which contrasts with the asynchronous, *event-driven* nature of sensors such as the *Dynamic Vision Sensor*. Introducing sequential, clock-based communication into a neuromorphic system would drastically degrade performance. As a result, it becomes essential to optimize both the efficiency and communication protocols beyond those adopted in traditional architectures [75]. To address this challenge, a dedicated communication protocol known as *Address Event Representation (AER)* has been proposed [76]. The functioning of the *AER* protocol is illustrated in Figure 14. In this scheme, an *address encoder* monitors all input lines for incoming events and performs multiplexing of these events onto a shared output bus. Through an arbitration mechanism, the encoder selects one event at a time based on its temporal order of arrival. It is important to note that the *AER* protocol does not provide an absolute timestamp for each event. Rather, the relative position of an event in the output sequence reflects the order in which it reached the encoder inputs. This allows the system to preserve event chronology without imposing the need for a global clock, thereby maintaining the asynchronous and parallel nature of neuromorphic communication.

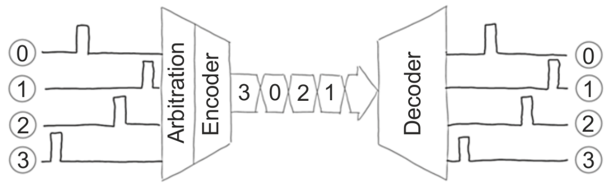


Figure 14: functioning of an AER protocol

PIXEL WORKING PRINCIPLE

As previously introduced, the pixels of a *Dynamic Vision Sensor* operate asynchronously and independently from one another. A simplified schematic of their internal structure is shown in Figure 15. This architecture comprises three main stages that work in cascade to convert changes in incident light into asynchronous events. The first stage is responsible for converting light intensity into a voltage signal that varies in proportion to the logarithmic change in irradiance. This is typically achieved using a photodiode, an amplifier, and a feedback *NMOS* transistor that ensures low-latency response and automatic gain control of the pixel. The voltage V_p across the photodiode depends logarithmically on the generated photocurrent, which is itself proportional to the local light intensity I. This relationship is described by:

$$V_p = V_{th} \ln\left(\frac{I}{I_0}\right), \tag{2.22}$$

where V_{th} denotes the thermal voltage and I_0 is the diode reverse saturation current.

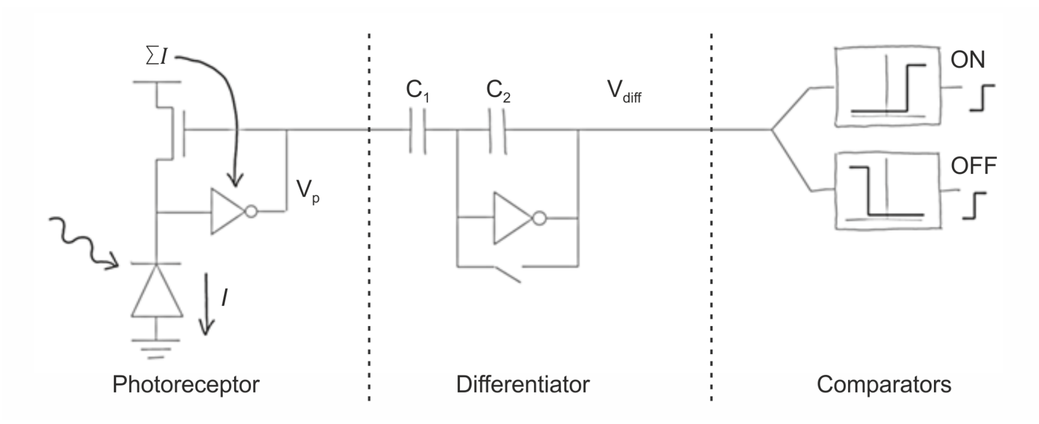


Figure 15: internal structure of a DVS pixel

The second stage generates a differential voltage that reflects the deviation of the current light intensity from the most recently stored value. Depending on the state of an internal switch, this block may either update the reference voltage used for future comparisons (when the switch is closed) or operate in differential mode (when the switch is open). In this second mode, the block behaves as an inverting amplifier with gain:

$$G = \frac{C_1}{C_2} \tag{2.23}$$

resulting in a voltage variation expressed as:

$$\Delta V_{diff} \cong -G\Delta V_p = -GV_{th}\Delta \ln\left(\frac{I}{I_0}\right). \tag{2.24}$$

This enables the sensor to continuously monitor variations in brightness relative to the previously observed level, rather than relying on an absolute static threshold. The third and final stage comprises two voltage comparators, one for positive and one for negative thresholds. When the variation in voltage exceeds a pre-defined contrast level (either positively or negatively) an event is generated in the form of a voltage spike. This event is then encoded and transmitted using the *AER* protocol.

2.7 CAMERA CALIBRATION

Camera calibration is the process of estimating its intrinsic parameters $\boldsymbol{K}$ and extrinsic parameters $[\boldsymbol{R} \quad | \quad \boldsymbol{t}]$ [71]. The first approach, known as calibration by resection, is based on the principle that if the image projections of $3D$ points with known coordinates (control points) are available, the unknown parameters can be determined by solving the perspective projection equations. In calibration by resection, a non-planar calibration object is required, whose control point coordinates must be measured with high precision. This requirement is challenging without precision instrumentation. Instead, it is often much easier to use a planar calibration target.

The *Sturm–Maybank–Zhang* (*SMZ*) calibration method [77,78] addresses this by requiring multiple (at least three) images of a single plane as shown in Figure 16, in contrast to resection, which needs only one image but of at least two distinct planes.

The relationship between a planar surface Π in the $3D$ scene and its perspective image can be described by a homography, which is a non-singular 3×3 matrix $\boldsymbol{H}$. By choosing the world coordinate system such that the plane Π lies at $Z = 0$, the perspective projection equation simplifies, showing that points on the plane are mapped to the image by $\boldsymbol{H}$. To estimate $\boldsymbol{H}$, a planar calibration target with a known pattern (e.g., a chessboard as shown in Figure 16) is used, allowing easy correspondence between known points on the plane and their projections in the image. The *Direct Linear Transform* (*DLT*) *algorithm* computes the homography from these correspondences.

Figure 16: three different views of a chessboard used for camera calibration

Since the camera projection matrix can be decomposed as Equation (2.13), the homography, considering $\boldsymbol{R} = [\boldsymbol{r_1} \quad \boldsymbol{r_2} \quad \boldsymbol{r_3}]$, can be expressed as:

$$\boldsymbol{H} = \lambda \boldsymbol{K}[\boldsymbol{r_1} \quad \boldsymbol{r_2} \quad \boldsymbol{t}], \tag{2.25}$$

where λ is an unknown scale factor.

Using the orthonormality constraints of the rotation matrix columns, one can derive equations that impose constraints on the intrinsic parameters $\boldsymbol{K}$. These constraints are typically written in terms of a symmetric matrix $\boldsymbol{B} = (\boldsymbol{K}\boldsymbol{K}^T)^{-1}$, which has six independent elements. By vectorizing the constraints, a linear system is formed whose solution yields the intrinsic parameters of the camera.

3

COMPUTER VISION FOR VIBRATIONS

This chapter provides an overview of key computer vision techniques employed for $3D$ flexural vibration reconstruction and analysis based on image sequences. The discussion begins with marker tracking methods and Digital Image Correlation (*DIC*), which are used to extract displacement information from image data by tracking local patterns or fiducial markers across multiple frames. The chapter then introduces the principles of epipolar geometry, which define the intrinsic geometric relationships between two views of the same scene and form the theoretical foundation for stereo vision. Building upon this framework, the chapter presents linear triangulation techniques, which allow for the estimation of three-dimensional point coordinates given multiple $2D$ projections and known camera parameters. Finally, an alternative approach based on frequency domain triangulation is explored. The methodologies and formulations introduced here are later extended and applied in Chapters 5 and 8, where they are used to reconstruct the vibration fields of different structural configurations using both conventional and event-based cameras.

3.1 MARKER TRACKING AND DIGITAL IMAGE CORRELATION METHODS

Information contained in consecutive $2D$ images must be extracted in order to analyse the temporal evolution of the object or of the image of a point $\boldsymbol{M} = [X, Y, Z]$. For this purpose, specific features are selected within the images to enable tracking the temporal evolution of $\boldsymbol{m} = [u, v]$ across frames. The main target-tracking methods can be categorized into point tracking and digital image correlation [79,80].

POINT TRACKING

The point-tracking approach is based on identifying and following distinct points of interest in the image sequence to estimate the target motion. It typically involves a limited number of features and provides precise information about their trajectories and positional variations over time. In this doctoral research, the positions of the markers are determined through a matching algorithm that seeks the highest similarity between a predefined template and the perspective-corrected image of the specimen. To facilitate the matching process, an ideal grid of nominal positions is introduced, where each node of the grid serves as the initial search point for the nearest marker. To compensate for perspective distortions introduced by the camera viewpoint, a planar homography transformation $\boldsymbol{H} \in \mathbb{R}^{3\times3}$ is computed, as shown in Figure 17. This transformation maps the original image $\boldsymbol{m_i}$ onto a rectified view improving detection robustness. The homography is estimated using known world coordinates $\boldsymbol{M_i}$ and their corresponding image projections $\boldsymbol{m_i}$, expressed in homogeneous coordinates. Within the rectified image, candidate marker centres are identified using template matching based on normalized cross-correlation between a predefined patch and the local image content.

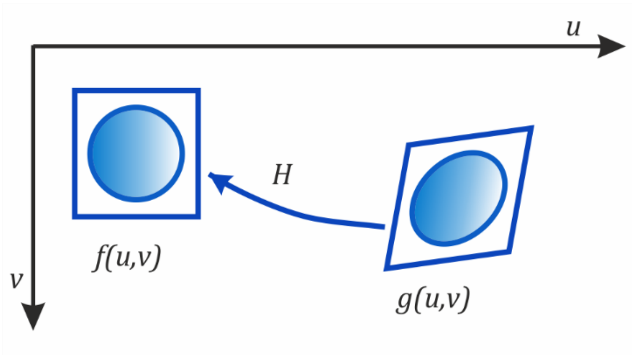


Figure 17: planar homography transformation to compensate camera distorsions

Due to imaging noise and varying illumination, false positives may occur. To robustly associate the detected markers with the theoretical grid positions, the problem is formulated as an assignment task. Given two sets: B containing the nominal marker coordinates and D containing the detected positions, and given a cost function $C(b,d)$ defined as the Euclidean distance, the optimal *one-to-one* correspondence $\psi : A \to B$ is found by minimizing the total assignment cost:

$$\min_{\psi} \sum_{b \epsilon B} C\big(b, \psi(b)\big) = \sum_{b \epsilon B} \|b - \psi(b)\|^2. \tag{3.1}$$

Note that set D can be larger than set B, therefore this is the unbalanced version of the assignment problem. This combinatorial optimization problem is solved efficiently using the *Kuhn–Munkres algorithm* [81], which guarantees an optimal solution in polynomial time. To improve localization accuracy, a subpixel refinement step is applied to each matched position. Around the peak of the correlation response, a one-dimensional parabolic fit is performed in both the horizontal and vertical directions using the central point and its immediate neighbours. The verices of the fitted parabolae provide a refined estimate of the marker position. Subsequently, the inverse homography H^{-1} is applied to map the refined marker coordinates back to the original, unrectified image space.

DIGITAL IMAGE CORRELATION

Digital image correlation (*DIC*) determines surface displacement and deformation by analysing pixel intensity patterns, such as brightness or greyscale, within a designated *DIC* window. By comparing these patterns across consecutive frames, *DIC* computes displacement fields and deformation maps. This method is particularly suited for capturing full-field measurements over extended surface regions. The principle underlying the *DIC* technique is based on the *Lucas–Kanade method* [82], originally developed for digital image alignment and optical flow estimation by exploiting variations in greyscale intensity values[83]. The primary goal of image registration procedures is to align the current image under analysis, denoted by $g(\boldsymbol{m}) = f(\boldsymbol{m}, t_{ref} + \Delta t)$, with a reference image $f(\boldsymbol{m}) = f(\boldsymbol{m}, t_{ref})$ via a chosen geometric transformation $\zeta(\boldsymbol{m}, \boldsymbol{A})$, with $\boldsymbol{A}$ indicating optimal parameters used to align the current image with the reference, as shown in Figure 18.

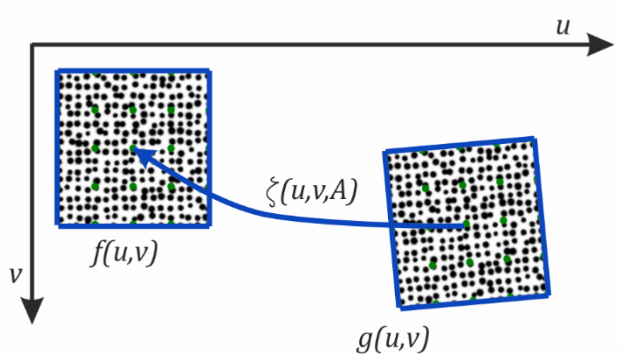


Figure 18: geometric transformation to align the current image with the reference

This approach formulates optical flow estimation as an optimization problem, where a selected region of the current image is iteratively aligned to the reference image. In its simplest form, the objective function minimized is the Sum of Squared Differences (*SSD*):

$$SSD = \sum_{\boldsymbol{m}} \left[g\big(\zeta(\boldsymbol{m}, \boldsymbol{A})\big) - f(\boldsymbol{m}) \right]^2. \tag{3.2}$$

By expanding Equation (3.2) in a Taylor series around an initial estimate of $\boldsymbol{A}$, it is obtained [83]:

$$SSD(\boldsymbol{m}, \boldsymbol{A} + \Delta\boldsymbol{A}) = \sum_{\boldsymbol{m}} \left[g(\boldsymbol{m} + \boldsymbol{A}) - \frac{\partial g}{\partial u} \Delta A_0 - \frac{\partial g}{\partial v} \Delta A_1 - f(\boldsymbol{m}) \right]^2. \tag{3.3}$$

In this formulation, $\boldsymbol{A}$ is the current estimate of the displacement vector, and ΔA_0, ΔA_1 are the parameter increments in each iteration. The optimal increments are found by setting the partial derivatives of Equation (3.3) with respect to $\Delta\boldsymbol{A}$ to zero. At each iteration, the solution for $\Delta\boldsymbol{A} = [\Delta A_0, \Delta A_1]^T$ is used to update the displacement estimation. The reference image region $f(\boldsymbol{m})$ is defined as a $l \times h$ window cantered around the observed point $\boldsymbol{m}$, where displacement is to be measured:

$$f(\boldsymbol{m}) = f\left(\left[u - \frac{l}{2}, \quad u + \frac{l}{2} \right], \quad \left[v - \frac{h}{2}, \quad v + \frac{h}{2} \right] \right). \tag{3.4}$$

This region is also referred to as the Region of Interest (*ROI*).

The transformed image region $g(\zeta(\boldsymbol{m}, \boldsymbol{A}))$ is the area in the current image that best aligns with the reference region according to the chosen optimization criterion. For the optimal increment $\Delta\boldsymbol{A}$ to be uniquely determined, the intensity pattern within the reference area $f(\boldsymbol{m})$ must be sufficiently distinctive in the vicinity of the current estimate $g(\zeta(\boldsymbol{m}, \boldsymbol{A}))$. Consequently, it is common practice to apply a random speckle pattern to the surface of observed objects to ensure uniqueness.

3.2 EPIPOLAR GEOMETRY

Epipolar geometry characterizes the intrinsic projective relationship between two images of the same three-dimensional scene captured from distinct viewpoints. As shown in Figure 19, given a point $\boldsymbol{m_1}$ in the first image, its corresponding point $\boldsymbol{m_2}$ in the second image is constrained to lie on a specific line, known as the epipolar line associated with $\boldsymbol{m_1}$. This constraint arises from the fact that $\boldsymbol{m_2}$ must be located on the intersection between the second image plane and the epipolar plane which is the plane defined by the original point $\boldsymbol{m_1}$ and the two camera centres $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$. The epipolar plane contains all possible projections of points lying along the optical ray passing through $\boldsymbol{m_1}$. All epipolar lines in an image converge at a unique point called the epipole, which corresponds to the projection of the opposing camera centre onto the image plane. The set of epipolar planes form a pencil of planes sharing the *baseline*, the line segment connecting the two camera centres $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$.

When the camera centre of projection lies in the focal plane of the other camera, the epipole is at infinity, and the epipolar lines form a family of parallel lines. A notable special case, known as the normal configuration, occurs when both epipoles lie at infinity; this happens when the baseline is contained within both focal planes, making the image planes parallel to the baseline. In this scenario, epipolar lines appear as parallel bundles in both images.

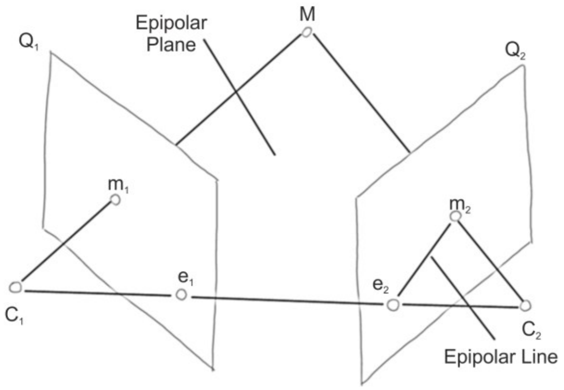


Figure 19: epipolar geometry representation

The bilinear relation linking corresponding points $\boldsymbol{m_1}$ and $\boldsymbol{m_2}$ is mathematically described by the fundamental matrix $\boldsymbol{F}$, which encodes all epipolar constraints. This matrix satisfies the *Longuet-Higgins equation*:

$$\boldsymbol{m_2^T F m_1} = 0 \tag{3.5}$$

expressing that the point $\boldsymbol{m_2}$ must lie on the epipolar line associated with $\boldsymbol{m_1}$. The fundamental matrix is of rank two, defined up to a scale factor, and contains seven degrees of freedom after accounting for intrinsic constraints.

3.3 LINEAR TRIANGULATION

After having identified the $2D$ coordinates in each view, the triangulation procedure allows to obtain the $3D$ coordinates of target points. It is also referred to as forward intersection and represents the process of reconstructing the $3D$ position of a point in object space from its corresponding projections in two images, given the respective camera projection matrices. Consider a $3D$ point $\boldsymbol{M}$ projected as $\boldsymbol{m_1}$ in the first image through the camera with projection matrix $\boldsymbol{P_1}$. The relationship between $\boldsymbol{m_1}$, $\boldsymbol{P_1}$, and $\boldsymbol{M}$ is expressed by the homogeneous equation:

$$[\boldsymbol{m_1}]_\times \boldsymbol{P_1 M} = 0, \tag{3.6}$$

where the cross-product matrix $[\boldsymbol{m_1}]_\times$ ensures the point $\boldsymbol{m_1}$ lies on the projection of $\boldsymbol{M}$. This equation provides two linearly independent constraints on the four homogeneous coordinates of $\boldsymbol{M}$. By similarly considering the homologous point $\boldsymbol{m_2}$ in the second image with projection matrix $\boldsymbol{P_2}$, the two sets of constraints can be combined into a single homogeneous linear system:

$$\begin{bmatrix}[\boldsymbol{m_1}]_\times \boldsymbol{P_1 M} \\ [\boldsymbol{m_2}]_\times \boldsymbol{P_2 M}\end{bmatrix} = \boldsymbol{0}, \tag{3.7}$$

which consists of six equations with four unknowns, of which four are linearly independent under ideal conditions. The solution $\boldsymbol{M}$ corresponds to the null space of this system and can be uniquely determined when the matrix has rank three. In practice, noise and measurement errors render the system inconsistent, thus the solution is approximated via a *least-squares method*, commonly implemented using singular value decomposition. This approach, known as the *linear-eigen method*, finds $\boldsymbol{M}$ as the right singular vector corresponding to the smallest singular value of the coefficient matrix. The two-camera configuration is illustrated in Figure 20a, where the area representing the possible location of the $3D$ point is elongated along the depth direction, reflecting the limited accuracy in this component when only two viewing rays are available.

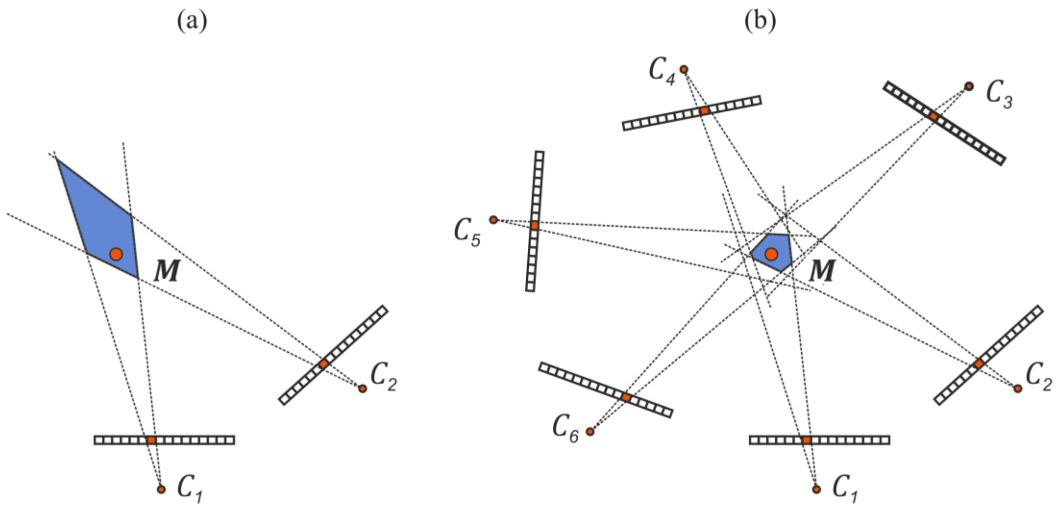


Figure 20: triangulation of M using a) two different camera views and b) six different camera views

The approach described can be extended to configurations with more than two cameras as shown in Figure 20b. Specifically, each additional camera contributes two independent constraints, resulting in a homogeneous linear system with $2N$ equations in four unknown variables when $N > 2$. It is important to note that a single camera yields only two independent equations, which are insufficient to recover the $3D$ coordinates of a point. This limitation underscores the fundamental requirement for at least two intersecting viewing rays to localize a point in space. As shown in Figure 20b, increasing the number of cameras progressively reduces the region of uncertainty in which the $3D$

point can be located, thereby improving positional accuracy, particularly along the depth direction.

3.4 FREQUENCY DOMAIN TRIANGULATION

After identifying the $2D$ displacement field in each view, for example using the *DIC* method [80,82] described in Section 3.1, the next step is to transform this information from the time domain to the frequency domain, thereby obtaining frequency-domain images [39]. A particular spatial point oscillating at frequency ω_k with a complex-valued spatial amplitude $\Delta\boldsymbol{M}(\omega_k)$ around an initial position $\boldsymbol{M}_{ref}$ is imaged by a perspective camera (see Equation 2.3). At the sampling time t_n the $3D$ position of the point is $\boldsymbol{M}(t_n)$, while the measured image-plane displacement $\Delta\boldsymbol{m}(t_n)$ represents the difference between its current and reference image positions:

$$\Delta\boldsymbol{m}(t_n) = \frac{1}{\lambda_n}\boldsymbol{PM}(t_n) - \frac{1}{\lambda_{ref}}\boldsymbol{PM}_{ref}, \tag{3.8}$$

where λ denotes the scale factors:

$$\begin{cases} \lambda_n = \boldsymbol{P}^{\{3\}}\boldsymbol{M}(t_n) \\ \lambda_{ref} = \boldsymbol{P}^{\{3\}}\boldsymbol{M}_{ref} \end{cases} \tag{3.9}$$

here, $\boldsymbol{P}^{\{3\}}$ represents the third row of $\boldsymbol{P}$.

Applying the discrete Fourier transform to both sides of Equation 3.8 across N images yields the frequency-domain amplitude of the image displacement at ω_k:

$$\Delta\widetilde{\boldsymbol{m}}(\omega_k) = \frac{1}{N}\sum_{\boldsymbol{n=1}}^{\boldsymbol{N}}\left(\frac{1}{\lambda_n}\boldsymbol{PM}(t_n) - \frac{1}{\lambda_{ref}}\boldsymbol{PM}_{ref}\right)e^{-2i\pi nk/N}. \tag{3.10}$$

Because the size of the observed object may not be negligible relative to its distance from the camera, perspective effects on the imaged point positions should generally be accounted for [71]. However, since the spatial position can be expressed as $\boldsymbol{M}(t_n) = \Delta\boldsymbol{M}(t_n) + \boldsymbol{M}_{ref}$ and the relative displacements $\Delta\boldsymbol{M}$ in high-frequency vibrations are typically orders of magnitude smaller, it is reasonable to assume:

$$\boldsymbol{P}^{\{3\}}\left(\Delta\boldsymbol{M}(t_n) + \boldsymbol{M}_{ref}\right) \cong \boldsymbol{P}^{\{3\}}\boldsymbol{M}_{ref}. \tag{3.11}$$

Consequently, the scale factors can be considered equal, simplifying Equation 3.10 to:

$$\Delta\widetilde{\boldsymbol{m}}(\omega_k) = \frac{1}{\lambda_{ref}}\boldsymbol{P}\frac{1}{N}\sum_{\boldsymbol{n=1}}^{\boldsymbol{N}}\left(\boldsymbol{M}(t_n) - \boldsymbol{M}_{ref}\right)e^{-2i\pi nk/N}$$
$$= \frac{1}{\lambda_{ref}}\boldsymbol{P}\Delta\boldsymbol{M}(\omega_k). \tag{3.12}$$

To reconstruct a frequency-domain image of the deflected point with correct perspective scaling, the relative displacement amplitudes $\Delta\widetilde{\boldsymbol{m}}(\omega_k)$ from Equation 3.12 must be mapped back to their reference image positions $\boldsymbol{m}_{ref}$, while the spatial coordinates $\Delta\boldsymbol{M}(\omega_k)$ and $\boldsymbol{M}_{ref}$ remain in homogeneous coordinates:

$$\widetilde{\boldsymbol{m}}(\omega_k) = \Delta\boldsymbol{M}(\omega_k) + \boldsymbol{M}_{ref} = \frac{1}{\lambda_{ref}}\boldsymbol{P}\left(\Delta\boldsymbol{M}(\omega_k) + \boldsymbol{M}_{ref}\right)$$
$$= \frac{1}{\lambda_{ref}}\boldsymbol{P}\boldsymbol{M}(\omega_k). \tag{3.13}$$

Once frequency-domain images have been computed, triangulation in the frequency domain can be performed. For each camera view, two homogeneous equations can be written for each image point $\boldsymbol{m}$. In an *n-view* setup, the resulting homogeneous system is solved using the least-squares method. After applying multiview triangulation across the entire frequency range, the amplitude spectrum of the point harmonic motion can finally be recovered:

$$\Delta\boldsymbol{M}(\omega_k) = \boldsymbol{M}(\omega_k) - \boldsymbol{M}_{ref} \quad \text{for} \quad k \in [1, N]. \tag{3.14}$$

4

FORMULATIONS FOR SOUND RADIATION

This chapter presents the theoretical foundations and numerical methodologies employed in the acoustic analysis of radiating structures which have been used in this PhD project to reconstruct the radiated sound field from the vibration field measured with optical cameras. It begins with the formulation of the wave equation in three dimensions and its expression in spherical coordinates under the assumption of isotropic media. The derivation of the inhomogeneous Helmholtz equation and its implications for time-harmonic problems are then discussed. Subsequently, the chapter introduces Green's functions and their role in the integral formulation of acoustic fields, leading to the Kirchhoff–Helmholtz integral equation. Given the limitations of analytical solutions for complex geometries, the Boundary Element Method (*BEM*) is introduced as a well-established numerical approach for solving exterior acoustic problems. Finally, attention is given to the computation of radiated sound power through time-averaged intensity, including its formulation for harmonic fields. This provides the necessary framework for evaluating acoustic performance based on pressure and particle velocity distributions. The theoretical and numerical concepts summarized in this chapter, although well established in the literature, are further developed and extended in Chapters 6 and 7, where they form the basis for the proposed formulations of sound power estimation and their application to increasingly complex structural configurations.

4.1 CONSERVATION LAWS

In the context of linear acoustics, the behaviour of small perturbations in a compressible, inviscid fluid can be described using the conservation laws of mass, momentum, and energy, along with appropriate thermodynamic relations [4,6,55,84]. Assuming a continuous and homogeneous medium, the differential form of the mass conservation equation is:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \boldsymbol{v}) = 0, \tag{4.1}$$

where ρ is the local fluid density, $\boldsymbol{v}$ is the particle velocity.

The momentum conservation equation, using the mass conservation reported in Equation (4.1), is expressed as:

$$\frac{\partial}{\partial t}(\rho \boldsymbol{v}) + \nabla \boldsymbol{\sigma} + \rho \boldsymbol{v} \nabla \boldsymbol{v} = \boldsymbol{f} \tag{4.2}$$

with $\boldsymbol{f}$ representing external body forces and $\boldsymbol{\sigma}$ the stress tensor which can be split in two terms and written as follows:

$$\boldsymbol{\sigma} = p\boldsymbol{I} - \boldsymbol{\tau}, \tag{4.3}$$

where p represents the pressure, $\boldsymbol{\tau}$ the viscous stress tensor, and $\boldsymbol{I}$ the identity tensor. However, for most acoustic applications, viscous effects are negligible, and the fluid is treated as inviscid.

The energy conservation equation, accounting for internal energy e_i, kinetic energy $e_k = |\boldsymbol{v}|^2/2$, and heat flux $\boldsymbol{q}$, is given by:

$$\frac{\partial}{\partial t}\rho(e_i + e_k) + \nabla[\rho \boldsymbol{v}(e_i + e_k)] = -\nabla \boldsymbol{q} - \nabla(p\boldsymbol{v}) + \nabla(\boldsymbol{\tau}\boldsymbol{v}) + \boldsymbol{f}\boldsymbol{v}. \tag{4.4}$$

If now, the thermodynamics law for reversible processes is introduced ($Tds = de_i + pd\rho^{-1}$, where T represents the temperature and s the entropy of the system), Equation (4.4) can be rewritten as:

$$\rho T\left(\frac{\partial s}{\partial t} + \boldsymbol{v}\nabla s\right) = -\nabla \boldsymbol{q} + \nabla(\boldsymbol{\tau}\boldsymbol{v}) - \boldsymbol{v}(\nabla \boldsymbol{\tau}). \tag{4.5}$$

In the absence of viscosity and thermal conduction, this simplifies to the isentropic flow condition:

$$\frac{\partial s}{\partial t} + \boldsymbol{v}\nabla s = 0. \tag{4.6}$$

A reversible process implies that all transformations within the fluid occur in a quasistatic way, ensuring that the system remains in thermodynamic equilibrium throughout the process.

If an ideal compressible fluid is now considered, the definition of the total derivative $D/Dt = \partial/\partial t + \boldsymbol{v}\nabla$ can be employed and Equation (4.2) for a non-viscous fluid with no flux can be expressed as:

$$\rho \frac{D\boldsymbol{v}}{Dt} = -\nabla p + \boldsymbol{f}. \tag{4.7}$$

Additionally for an isentropic flow holds:

$$\frac{Ds}{Dt} = 0 \tag{4.8}$$

and combining Equation (4.8) with the equation of state, and the mass conservation (4.1), it is obtained:

$$\frac{Dp}{Dt} + \rho c^2 \nabla \boldsymbol{v} = 0, \tag{4.9}$$

where c is the speed of sound in the medium. Finally, by coupling the equations that govern the relationship between pressure and particle velocity, it is possible to derive the classic acoustic wave equation, as will be presented in the following Section 4.2.

4.2 ACOUSTIC WAVE EQUATION

Having established the conservation laws of mass and momentum, given the spherical symmetry of three-dimensional waves, and assuming that the wave propagates through an isotropic medium, then the spatial operator on the pressure in the three-dimensional wave equation can be written in spherical coordinates as:

$$\nabla^2 p = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial p}{\partial r}\right) \tag{4.10}$$

and the wave equation is:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial p}{\partial r}\right) - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0. \tag{4.11}$$

The solution of Equation (4.11) can be written in the form:

$$p(r,t) = \frac{f\left(t - \frac{r}{c}\right)}{r} + \frac{g\left(t + \frac{r}{c}\right)}{r} \tag{4.12}$$

where $f\left(t - \frac{r}{c}\right)/r$ represents pressure fluctuations which travel outwards from the origin of the spherical coordinates, while $g\left(t + \frac{r}{c}\right)/r$ represents pressure fluctuations which travel inwards towards the origin. For an unbounded medium, the Sommerfeld radiation condition requires:

$$\lim_{r\to\infty}\left\{r\left(\frac{\partial p}{\partial r} + \frac{1}{c}\frac{\partial p}{\partial t}\right)\right\} = 0, \tag{4.13}$$

which is satisfied only by $f\left(t-\frac{r}{c}\right)/r$.

Once the homogeneous wave equation (Equation (4.11)) has been derived, it becomes necessary to account for the presence of structures within the acoustic domain. This is achieved by specifying appropriate boundary conditions, which describe the physical constraints imposed at the surfaces of the structure and at the limits of the domain. Given that the wave equation is second order, only one boundary condition is required at each point on the boundary, which may be defined in terms of pressure, particle velocity, or a combination of both. The boundary surface Ω (see Figure 21) is typically partitioned into three distinct, non-overlapping regions, depending on the type of boundary condition applied:

$$\Omega = \Omega_{\mathrm{D}} + \Omega_{\mathrm{N}} + \Omega_{R}. \tag{4.14}$$

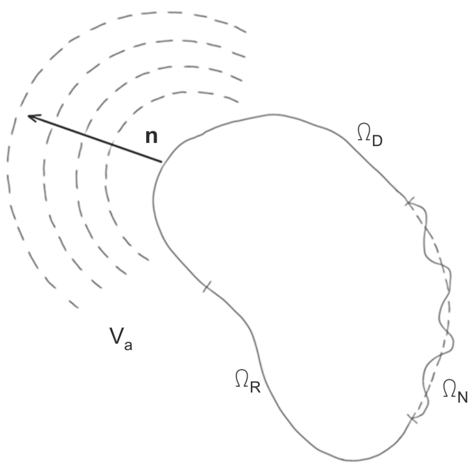


Figure 21: Exterior vibroacoustic problem boundary conditions

In fact, it is possible to have:

- Dirichlet boundary condition Ω_{D} specifies the acoustic pressure directly:

$$p(t) = p_0(t) \qquad \text{on } \Omega_{\mathrm{D}}. \tag{4.15}$$

- Neumann boundary condition Ω_{N} defines the normal component of the particle velocity:

$$v_n(t) = v_{n0}(t) \qquad \text{on } \Omega_{\mathrm{N}}. \tag{4.16}$$

- Robin boundary condition Ω_R introduces a linear relation between pressure and normal velocity through a local acoustic impedance Z:

$$v_n(t) = \frac{1}{Z} p(t) \qquad \text{on } \Omega_R. \tag{4.17}$$

The Robin condition generalizes the previous two cases: as $\boldsymbol{Z} \rightarrow \boldsymbol{0}$, it reduces to a Dirichlet condition; as $\boldsymbol{Z} \rightarrow \infty$, it approximates a Neumann condition.

Considering an acoustic system being in its steady state with a time harmonic excitation, the acoustic pressure and velocity can also be written as time harmonic expressions. More specifically, the sound pressure can be written as:

$$p(\boldsymbol{x_a}, t) = \mathcal{R}e\{p(\boldsymbol{x_a}, \omega) e^{i\omega t}\}, \tag{4.18}$$

where $\boldsymbol{x_a}$ is a position in the medium. Therefore, the definition of the wave equation can be written as:

$$\nabla^2 p(\boldsymbol{x_a}) + k^2 p(\boldsymbol{x_a}) = \boldsymbol{f}(\boldsymbol{x_a}). \tag{4.19}$$

Here $k = \omega / c$ represents the wavenumber. Equation (4.19) represents the inhomogeneous form of the classic Helmholtz differential equation. Since it involves only spatial derivatives, it constitutes a Partial Differential Equation (*PDE*) that depends solely on the spatial coordinates within the acoustic domain. However, since each wavenumber k, and therefore each frequency, leads to a different *PDE*, the Helmholtz equation must be solved separately for each k within the frequency range $[\omega\_\min, \omega_{max}]$. Additionally, by assuming time-harmonic behaviour of the acoustic field, a direct relationship between the particle velocity $\boldsymbol{v}$ and the pressure gradient ∇p can be derived:

$$\nabla p = i\omega\rho\boldsymbol{v}. \tag{4.20}$$

4.3 KIRCHHOFF-HELMHOLTZ EQUATION

In acoustics, the assumption of linearity is often adopted to simplify analysis. Under linearity, the system response scales proportionally with the applied load and follows the principle of superposition: *the total response to multiple sources equals the sum of individual responses*. This allows for a complete characterization of the system via its unit response, typically expressed as a *FRF* $\mathcal{H}(\boldsymbol{x_a}, \boldsymbol{x}, \omega)$ where $\boldsymbol{x_a}$ represents the position at which the acoustic response is evaluated and $\boldsymbol{x}$ represents the location of a unit source. To determine $\mathcal{H}$ over the entire frequency range, the system can be excited at $\boldsymbol{x}$ by a *Dirac impulse* $\delta(\boldsymbol{x}, t)$ in the time domain, which corresponds to uniform excitation in the frequency domain. Under such excitation, the inhomogeneous Helmholtz Equation (4.19) becomes:

$$\nabla^2 p(\boldsymbol{x_a}) + k^2 p(\boldsymbol{x_a}) = -4\pi\delta(\boldsymbol{x}) \tag{4.21}$$

and the solution of this differential equation can be split in a homogeneous p_{hom} and a particular form p_{par} [85]:

$$\begin{cases} p_{hom}(\boldsymbol{x_a}) = -i\dfrac{\sin(k\boldsymbol{r})}{4\pi\boldsymbol{r}} \\ \\ p_{par}(\boldsymbol{x_a}) = \dfrac{\cos(k\boldsymbol{r})}{4\pi\boldsymbol{r}} \end{cases} \tag{4.22}$$

where $\boldsymbol{r} = |\boldsymbol{x_a} - \boldsymbol{x}|$. Combining both gives the total solution, or acoustic Green's function G in a three-dimensional space:

$$G(\boldsymbol{x_a}, \boldsymbol{x}) = \frac{e^{-ik\boldsymbol{r}}}{4\pi\boldsymbol{r}}. \tag{4.23}$$

This Green's function describes the propagation of a spherical wave from a point source and inherently satisfies the Sommerfeld radiation conditions at infinity (Equation (4.13)), ensuring physical consistency in unbounded domains. It is important to highlight that the Green's function is inherently a complex-valued function. It exhibits a singularity as the distance $\boldsymbol{r} = |\boldsymbol{x_a} - \boldsymbol{x}|$ approaches zero. Moreover, the function is frequency-dependent, since the acoustic wavenumber k varies with the angular frequency ω, thereby influencing the spatial behaviour of the solution.

To determine the radiated acoustic field from objects with irregular geometries, the Kirchhoff–Helmholtz integral equation is employed. Its derivation relies on Green's second identity, which states that for any two sufficiently smooth and non-singular scalar functions φ and ψ defined in a volume V bounded by a surface Ω_s, the following holds:

$$\int_{\Omega_s} \left(\varphi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\varphi}{\partial n}\right) d\Omega_s = \int_V (\varphi\nabla^2\psi - \psi\nabla^2\varphi)\, dV. \tag{4.24}$$

In the acoustic context, the function φ is replaced by the acoustic pressure $p(\boldsymbol{x_a})$, while ψ corresponds to the free-space Green's function $G(\boldsymbol{x_a}, \boldsymbol{x})$. Both functions satisfy the homogeneous Helmholtz equation within the acoustic domain V_a, which is enclosed by the radiating surface Ω and two concentric spheres: a very small one Ω_a centred at the field point $\boldsymbol{x_a}$ and a much larger one Ω_∞ with radius tending to infinity. This configuration allows the application of the Sommerfeld radiation condition and avoids the singularity at $\boldsymbol{x_a} = \boldsymbol{x}$.

Substituting into the homogeneous Helmholtz equation the free-space Green's function $G(\boldsymbol{x_a}, \boldsymbol{x})$ (Equation (4.23)), and the acoustic pressure $p(\boldsymbol{x_a})$, it is obtained:

$$\int_{\Omega+\Omega_a+\Omega_\infty} \left(p(\boldsymbol{x_a}) \frac{\partial G(\boldsymbol{x_a}, \boldsymbol{x})}{\partial n} - G(\boldsymbol{x_a}, \boldsymbol{x}) \frac{\partial p(\boldsymbol{x_a})}{\partial n} \right) d\Omega = 0 \tag{4.25}$$

and the volume integral in Equation (4.24) vanishes because both $p(\boldsymbol{x_a})$ and $G(\boldsymbol{x_a}, \boldsymbol{x})$ satisfy the Helmholtz equation. The behaviour of the surface integrals depends on the location of the field point $\boldsymbol{x_a}$, resulting in three distinct cases:

- When $\boldsymbol{x_a}$ is inside V_a: the integral over Ω_a converges to $p(\boldsymbol{x})$, and the integral over Ω_∞ vanishes due to the radiation condition. The integral reduces to:

$$\int_{\Omega} \left(p(\boldsymbol{x_a}) \frac{\partial G(\boldsymbol{x_a}, \boldsymbol{x})}{\partial n} - G(\boldsymbol{x_a}, \boldsymbol{x}) \frac{\partial p(\boldsymbol{x_a})}{\partial n} \right) d\Omega = -p(\boldsymbol{x}). \tag{4.26}$$

- When $\boldsymbol{x_a}$ lies on the surface Ω: assuming the surface is smooth and the normal direction is well-defined, the integral over Ω_a converges to half the field value:

$$\int_{\Omega} \left(p(\boldsymbol{x_a}) \frac{\partial G(\boldsymbol{x_a}, \boldsymbol{x})}{\partial n} - G(\boldsymbol{x_a}, \boldsymbol{x}) \frac{\partial p(\boldsymbol{x_a})}{\partial n} \right) d\Omega = -\frac{1}{2} p(\boldsymbol{x}). \tag{4.27}$$

- When $\boldsymbol{x_a}$ is outside V_a: both surface integrals over Ω_a and Ω_∞ vanish, and the total integral is zero:

$$\int_{\Omega} \left(p(\boldsymbol{x_a}) \frac{\partial G(\boldsymbol{x_a}, \boldsymbol{x})}{\partial n} - G(\boldsymbol{x_a}, \boldsymbol{x}) \frac{\partial p(\boldsymbol{x_a})}{\partial n} \right) d\Omega = 0. \tag{4.28}$$

These three situations are conveniently expressed in the Kirchhoff–Helmholtz integral equation [58], which generalizes the acoustic pressure field $p(\boldsymbol{x_a})$ at any point $\boldsymbol{x_a}$ as:

$$C(\boldsymbol{x}) p(\boldsymbol{x}) = \int_{\Omega} \left(p(\boldsymbol{x_a}) \frac{\partial G(\boldsymbol{x_a}, \boldsymbol{x})}{\partial n} + i\rho\omega G(\boldsymbol{x_a}, \boldsymbol{x}) v_n(\boldsymbol{x_a}) \right) d\Omega. \tag{4.29}$$

Here, $C(\boldsymbol{x})$ takes the values of:

$$\begin{cases} C(\boldsymbol{x}) = -1 & & \boldsymbol{x} \in V_a \\ C(\boldsymbol{x}) = 0 & \text{if} & \boldsymbol{x} \notin V_a \\ C(\boldsymbol{x}) = -1/2 & & \boldsymbol{x} \in \Omega \end{cases} \tag{4.30}$$

This formulation enables the evaluation of the pressure field throughout the domain from known boundary data on the radiating surface.

In the following sections, it is assumed that the boundary represents a vibrating surface. As a result, the particle velocity of the fluid at the boundary coincides with the surface velocity itself, therefore $v_n(\boldsymbol{x_a}) = \dot{u}_r(\boldsymbol{x})$.

4.4 BOUNDARY ELEMENT METHOD (*BEM*) IN ACOUSTICS

While the Kirchhoff-Helmholtz integral equation provides an exact analytical solution for simple geometries, such as spheres or cylinders with infinitely long baffles, its analytical evaluation becomes intractable for complex or irregular shapes. In such cases, the Kirchhoff-Helmholtz integral (Equation (4.29)) must be solved numerically using for example algorithms such as *finite element method* (*FEM*) or *boundary element method* (*BEM)*. This latter represents one of the most established numerical approaches for this purpose which is extensively employed in acoustic simulations. The method is typically categorized into *direct* and *indirect* formulations and both approaches follow a *two-step* procedure where the 'boundary variables' are first evaluated at the discrete set of points and then the 'field variable', that is, the sound pressure in the fluid domain, is obtained from the direct integral formulation, given in Equation (4.29). The *direct* approach imposes physical quantities, such as pressure and normal fluid velocity, on the boundary, and is generally limited to closed geometries, with some exceptions in semi-infinite configurations like tunnels. The *indirect* formulation, in contrast, offers greater flexibility by enabling the treatment of both open and closed geometries. However, it defines boundary conditions in terms of differences across boundary faces, complicating the physical interpretation of the resulting fields. Both formulations can be implemented using *collocation* or *Galerkin* schemes, depending on the desired numerical properties and accuracy. A major limitation of classic *BEM* lies in the nature of the resulting system matrices, which are fully populated, frequency-dependent, and typically lack desirable properties such as symmetry or positive definiteness. Consequently, solving large-scale problems becomes computationally demanding. In this thesis the *direct BEM* approach based on *collocation* methods has been employed, and on this method will focus the following discussion.

DIRECT COLLOCATIONAL *BEM*

In the first step of the Boundary Element Method, the surface of the radiating or scattering object is discretized into a set of small subdomains known as boundary elements. For acoustic applications, this is typically achieved using triangular or quadrangular elements. The boundary variables, namely the acoustic pressure p_s and the normal particle velocity $\dot{u}_r$, are approximated by shape functions defined locally over these elements. The surface Ω is thus partitioned into elements Ω_{e}, connected via nodal points generally placed at the element vertices. The boundary quantities are then interpolated using prescribed shape functions, each associated with a specific node, which take the value of one at their corresponding node and zero at all others. These functions are nonzero only over the subset of elements surrounding the given node.

Although constant shape functions (uniform over the element) could in principle be used, since the boundary integral equations do not involve spatial derivatives of p_s or

$\dot{u}_r$, they are rarely adopted due to poor convergence properties. Linear shape functions, which provide a piecewise-linear interpolation across each element, offer a more efficient and accurate representation and are therefore commonly employed. In practice, the approximation of the boundary variables with such local shape functions transforms the boundary integral formulation into a discrete system of equations, where the unknowns are the nodal values of pressure p_s and normal velocity $\dot{u}_r$. The global solution is assembled from contributions of each element, relying on local definitions of the shape functions rather than on global basis functions.

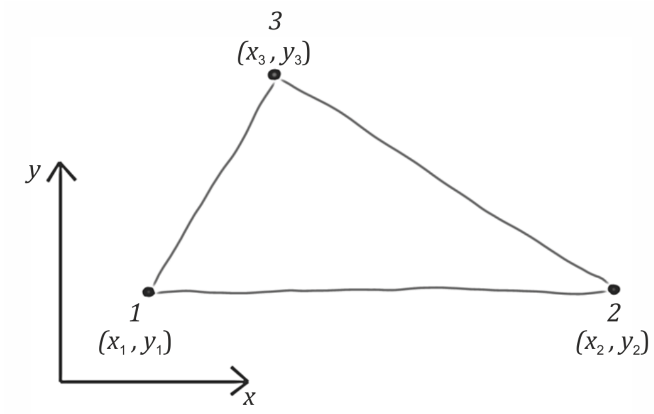


Figure 22: nodes of a generic triangular element

Let $\boldsymbol{x_n} = (x_n, y_n)$ with $n = 1,2,3$, be the coordinates of the three nodes of a generic triangular element shown in Figure 22. The pressure $p_s(\boldsymbol{x})$ and the normal velocity $\dot{u}_r(\boldsymbol{x})$ within the element are then approximated as:

$$\begin{cases} p_s(\boldsymbol{x}) = \phi_1 p_{s1} + \phi_2 p_{s2} + \phi_3 p_{s3} \\ \dot{u}_r(\boldsymbol{x}) = \phi_1 \dot{u}_{r1} + \phi_2 \dot{u}_{r2} + \phi_3 \dot{u}_{r3} \end{cases} \tag{4.31}$$

and in a more compact form:

$$\begin{cases} p_s(\boldsymbol{x}) = \boldsymbol{\phi} \begin{bmatrix} p_{s1} \\ p_{s2} \\ p_{s3} \end{bmatrix} \\ \dot{u}_r(\boldsymbol{x}) = \boldsymbol{\phi} \begin{bmatrix} \dot{u}_{r1} \\ \dot{u}_{r2} \\ \dot{u}_{r3} \end{bmatrix} \end{cases} \tag{4.32}$$

where p_{sn} and $\dot{u}_{rn}$ represent the values of pressure and normal velocity at node n, and ϕ_n are the shape functions defined by:

$$\phi_n = \frac{1}{2A}(A_n^0 + a_n x + b_n y). \tag{4.33}$$

The constants a_n, b_n and A_n^0 are determined from the nodal coordinates using the inverse of $\boldsymbol{D}$:

$$[\boldsymbol{D}]^{-1} = \frac{1}{2A}\begin{bmatrix} A_1^0 & A_2^0 & A_3^0 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}, \tag{4.34}$$

while A is the area of the triangular element, computed as:

$$A = \frac{1}{2}\det\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}. \tag{4.35}$$

Alternatively, the area can also be expressed using the coefficients a_n and b_n as:

$$A = \frac{1}{2}(a_1 b_2 - a_2 b_1). \tag{4.36}$$

Since both p_s and $\dot{u}_r$ are scalar fields without spatial derivatives in the boundary integral formulation, their interpolation via linear shape functions is both sufficient and efficient. These shape functions are defined locally on each element and, once assembled, guarantee global continuity across the discretized boundary, which is consistent with the *collocation-based BEM* approach adopted in many acoustic problems. This formulation allows the transformation of the continuous boundary integral equations into a discrete linear system, where the unknowns are the nodal values of pressure and/or normal velocity. The resulting approximation is valid across the discretized boundary and forms the basis for solving scattering and radiation problems using the Boundary Element Method.

At this point, after having defined the interpolation of the acoustic pressure p_s and the normal particle velocity $\dot{u}_r$ on each triangular element using shape functions $\boldsymbol{\phi}$, the continuous boundary integral formulation, specifically the Kirchhoff–Helmholtz integral (Equation (4.29)), can be discretised. The pressure and normal velocity fields are interpolated as:

$$\begin{cases} p_s(\boldsymbol{x}) \cong \sum_{n=1}^{N_n} p_{sn}\phi_n \\ \dot{u}_r(\boldsymbol{x}) \cong \sum_{n=1}^{N_n} \dot{u}_{rn}\phi_n \end{cases} \tag{4.37}$$

where N_n are the total number of boundary nodes. Substituting Equation (4.37) into the Kirchhoff-Helmholtz integral (Equation (4.29)) and evaluating it at each collocation point $\boldsymbol{x}$, it is obtained:

$$\begin{aligned} C(\boldsymbol{x})p_s(\boldsymbol{x}) + \sum_{n=1}^{N_n} \int_{\Omega_\mathrm{e}} p_{sn} \left(\frac{\partial G(\boldsymbol{x_a}, \boldsymbol{x})}{\partial n} \phi_n \right) d\Omega_\mathrm{e} \\ = i\omega\rho \sum_{n=1}^{N_n} \int_{\Omega_\mathrm{e}} \dot{u}_{rn} (G(\boldsymbol{x_a}, \boldsymbol{x}) \phi_n) d\Omega_\mathrm{e}. \end{aligned} \tag{4.38}$$

This system of equations can be rewritten in compact matrix form as:

$$[\boldsymbol{A}]\boldsymbol{p_s} = i\omega\rho[\boldsymbol{B}]\dot{\boldsymbol{u}}_{\boldsymbol{r}} + \boldsymbol{p_q}, \tag{4.39}$$

where $\boldsymbol{p_s}$ and $\dot{\boldsymbol{u}}_{\boldsymbol{r}}$ are the vectors of nodal pressures and normal velocities and $\boldsymbol{p_q}$ accounts for incident pressure if present:

$$\boldsymbol{p_s} = \begin{bmatrix} p_{s1} \\ p_{s2} \\ \vdots \\ p_{sNodes} \end{bmatrix}, \qquad \dot{\boldsymbol{u}}_{\boldsymbol{r}} = \begin{bmatrix} \dot{u}_{r1} \\ \dot{u}_{r2} \\ \vdots \\ \dot{u}_{rNodes} \end{bmatrix}, \qquad \boldsymbol{p_q} = \begin{bmatrix} p_{q1} \\ p_{q2} \\ \vdots \\ p_{qNodes} \end{bmatrix}. \tag{4.40}$$

Moreover, the matrices $[\boldsymbol{A}]$ and $[\boldsymbol{B}]$ have elements:

$$\begin{cases} A_{cn} = C(\boldsymbol{x})\partial_{cn} + \displaystyle\int_{\Omega_e} \frac{\partial G(\boldsymbol{x_a}, \boldsymbol{x})}{\partial n} \phi_n d\Omega_e \\ \\ B_{cn} = \displaystyle\int_{\Omega_\mathrm{e}} (G(\boldsymbol{x_a}, \boldsymbol{x}) \phi_n) d\Omega_\mathrm{e} \end{cases} \tag{4.41}$$

Here, c denotes the collocation point considered and ∂_{cn} is the *Kronecker delta*, defined as:

$$\partial_{cn} = \begin{cases} 1 \\ 0 \end{cases} \quad \text{if} \quad \begin{matrix} c = n \\ c \neq n \end{matrix} \tag{4.42}$$

These integrals are numerically evaluated over each triangular element using Gaussian quadrature or other appropriate integration schemes. After having derived the boundary surface results from the first step, the 'field variable', that is, the sound pressure in the fluid domain, is obtained.

4.5 TOTAL POWER RADIATED

The total sound radiation emitted by a vibrating structure can be conveniently quantified in terms of the time-averaged total radiated acoustic power, which is defined as:

$$\bar{P} = \lim_{T\to\infty} \frac{1}{T}\int_0^T P(t)dt. \tag{4.43}$$

Here, $P(t)$ denotes the instantaneous total radiated sound power at time t. Normally, this quantity is evaluated by integrating the acoustic intensity vector $\boldsymbol{I}(\boldsymbol{x_a}, t)$ over a closed surface Ω_a that surrounds the vibrating structure:

$$\boldsymbol{I}(\boldsymbol{x_a}, t) = p(\boldsymbol{x_a}, t)\boldsymbol{v}(\boldsymbol{x_a}, t), \tag{4.44}$$

where $p(\boldsymbol{x_a}, t)$ is the acoustic pressure at the position $\boldsymbol{x_a}$ and $\boldsymbol{v}(\boldsymbol{x_a}, t)$ is the particle velocity vector at the same location. In general, the velocity vector $\boldsymbol{v}(\boldsymbol{x_a}, t)$ (and thus the intensity) has three spatial components. However, if the surface Ω_a is sufficiently far from the radiating object, specifically, when its principal dimension is much larger than both the principal dimension of the object and the acoustic wavelength λ, the wavefront over Ω_a can be assumed to be locally planar. Under this assumption, the particle velocity is directed radially/normally outward, and the intensity vector simplified to:

$$\boldsymbol{I}(\boldsymbol{x_a}, t) \cong I_n(\boldsymbol{x_a}, t) = p(\boldsymbol{x_a}, t)v_n(\boldsymbol{x_a}, t), \tag{4.44}$$

where $v_n(\boldsymbol{x_a}, t)$ is the normal component of the particle velocity.

Consequently, the instantaneous radiated power becomes:

$$P(t) = \int_{\Omega_a} I_n(\boldsymbol{x_a}, t)d\Omega_a = \int_{\Omega_a} p(\boldsymbol{x_a}, t)v_n(\boldsymbol{x_a}, t)d\Omega_a. \tag{4.45}$$

This formulation provides a practical means to compute the total radiated acoustic energy by evaluating the surface integral over a measurement surface sufficiently distant from the source, under the assumption of locally plane wave propagation. Moreover, the normal component of the particle velocity, in the far field can be assumed to be:

$$v_n(\boldsymbol{x_a}, t) = \frac{p(\boldsymbol{x_a}, t)}{\rho c}. \tag{4.46}$$

Now, if time-harmonic fields for both pressure $p(\boldsymbol{x_a}, t) = \mathcal{R}e\{p(\boldsymbol{x_a}, \omega)e^{i\omega t}\}$ and particle velocity $v_n(\boldsymbol{x_a}, t) = \mathcal{R}e\{v_n(\boldsymbol{x_a}, \omega)e^{i\omega t}\}$ are considered, it is useful to express the time averaged acoustic intensity directly in terms of complex quantities. In this case, the instantaneous acoustic intensity is the product of these real-valued fields. Therefore, the time averaged acoustic intensity is given by:

$$I_n(\boldsymbol{x_a},\omega) = \frac{1}{2}[\mathcal{R}e\{p(\boldsymbol{x_a},\omega)\}\mathcal{R}e\{v_n(\boldsymbol{x_a},\omega)\}] \\ +\frac{1}{2}[\mathcal{I}m\{p(\boldsymbol{x_a},\omega)\}\mathcal{I}m\{v_n(\boldsymbol{x_a},\omega)\}], \quad (4.47)$$

which in a more compact form and using complex notation is:

$$I_n(\boldsymbol{x_a},\omega) = \frac{1}{2}[\mathcal{R}e\{p(\boldsymbol{x_a},\omega)v_n^*(\boldsymbol{x_a},\omega)\}] \\ = \frac{1}{2}[\mathcal{R}e\{p^*(\boldsymbol{x_a},\omega)v_n(\boldsymbol{x_a},\omega)\}], \quad (4.48)$$

where $p^*(\boldsymbol{x_a},\omega)$ and $v_n^*(\boldsymbol{x_a},\omega)$ are the complex conjugates of $p(\boldsymbol{x_a},\omega)$ and $v_n(\boldsymbol{x_a},\omega)$ respectively.

Equation (4.48) can be employed to obtain the sound power radiated in complex notation by substituting it into (4.45):

$$P(\omega) = \int_{\Omega_a} I_n(\boldsymbol{x_a},\omega)d\Omega_a = \frac{1}{2}\int_{\Omega_a} \mathcal{R}e\{p^*(\boldsymbol{x_a},\omega)v_n(\boldsymbol{x_a},\omega)\}d\Omega_a \quad (4.49)$$

and considering a surface Ω_a sufficiently distant from the source, the sound waves locally resemble a plane wave over the measurement surface Ω_a such that the particle velocity has radial direction given by Equation (4.46) and the instantaneous total sound power radiation considering time-harmonic sound waves results:

$$P(\omega) = \frac{1}{2\rho c}\int_{\Omega_a} |p(\boldsymbol{x_a},\omega)|^2 \mathrm{d}\Omega_a, \quad (4.50)$$

where $p(\boldsymbol{x}_a,\omega)$ is the complex amplitude of the time-harmonic sound pressure at position $\boldsymbol{x}_a$ on the measurement surface Ω_a.

Due to the very large range of amplitudes which are possible in acoustics, it is conventional to adopt a logarithmic measurement scale. This approach not only allows for a more practical representation of acoustic quantities but also aligns well with human auditory perception, which is approximately logarithmic in nature. Specifically, a doubling of the pressure amplitude is typically perceived as an equivalent increment in loudness. The standard logarithmic unit employed is the decibel scale (dB). In general terms, a level expressed in decibels is defined as ten times the *base-10* logarithm of the ratio between an energy-like quantity and a reference one:

$$Level\ (dB) = 10\log_{10}\left(\frac{quantity}{reference\ value}\right). \tag{4.51}$$

This formulation is broadly applicable to a variety of acoustic variables. The use of a logarithmic scale necessitates the specification of a reference quantity: a level of $0\ dB$ corresponds to a quantity equal to the reference, while negative decibel values indicate quantities smaller than the reference.

When this definition is applied to sound pressure, the relevant metric is the sound pressure level (*SPL*) which, according to Equation (4.51), it is expressed as:

$$SPL = 10\log_{10}\left(\frac{p^2}{p_{ref}^2}\right), \tag{4.52}$$

where p is the sound pressure and p_{ref} the reference value, which for sound in air is equal to $20\ \mu Pa$. This reference value corresponds to a pressure of a $1\ kHz$ pure tone at the threshold of hearing for an average young adult. Thus, this threshold of audibility defines the zero level of the sound pressure scale $SPL = 0\ dB$.

Similar to sound pressure, the sound power emitted by sources also spans several orders of magnitude. For instance, the power output of a whisper may be as low as $10^{-9}W$, whereas that of a jet engine can reach up to $10^{4}W$. To represent such a broad dynamic range, a logarithmic scale is likewise adopted for acoustic power. Accordingly to Equation (4.51), the Sound Power Level (*SWL*) is defined as:

$$SWL = 10\log_{10}\left(\frac{P}{P_{ref}}\right), \tag{4.53}$$

where P is the sound power of the source and $P_{ref} = 10^{-12}\ W$ is the standard reference sound power.

5

VIBRATION MEASUREMETS USING A SINGLE HIGH-SPEED CAMERA

This chapter presents the experimental framework developed to validate the frequency-domain triangulation approach for the reconstruction of three-dimensional vibration fields from single camera multi-view acquisitions. Two test structures are considered: a baffled cylindrical shell [50,86,87] and a small box, representative of thin-walled components with different geometric and modal characteristics. The experimental setups involve the use of a high-speed camera and a scanning laser vibrometer for full field flexural vibration acquisition. The application of the frequency-domain triangulation described in Chapter 3 is here applied at the two case studies. Moreover, an overview of the structural dynamics of cylindrical shells and thin plates is presented, focusing on their flexural natural modes. Finally, the reconstructed flexural vibration fields are analysed in terms of deflection shapes and time-averaged kinetic energy spectra, enabling a quantitative and qualitative comparison between high-speed camera and laser-based measurements.

5.1 TEST STRUCTURES: BAFFLED CYLINDER AND SMALL BOX

The vibration and sound radiation optical measurement techniques presented in this thesis are discussed with respect to model-structures representative of the typical applications encountered in practice. For this purpose, two different closed thin-shell structures were designed and built, as schematically shown in Figure 23.

The first structure is a baffled cylinder (Figure 23a), made of steel. The choice of a cylindrical geometry was motivated by its suitability for analytical solutions in the study of the acoustic radiation, as will be discussed in Chapter 6. The measurement cylinder has the dimensions listed in Table 2, and was equipped with a grid of 62×20 target points at which the Frequency Response Functions (*FRFs*) of the radial displacements per unit force excitation were measured. These points are located at the centres of the quadrilateral mesh elements used for the derivation of sound radiation. The cylinder is clamped on rigid flanges at both ends and is equipped with two cylindrical extensions acting as acoustically rigid baffles. The bending excitation is performed using a shaker fixed to the basement inside the cylinder and connected to the surface via a *stinger*. The structure is also equipped with an impedance head to record force measurements from the excitation and acceleration data.

Table 2: baffled cylinder general properties

BAFFLED CYLINDER: General Properties		
Thickness (mm)	1	
Radius (mm)	149	
Height (mm)	296	
N° target points	62×20	
Material cylinder and baffles	Steel	
	Density	$\rho_c = 7500\ kg/m^3$
	Young modulus	$E_c = 21E^{10}\ N/m^2$
	Poisson ratio	$\nu_c = 0.29$

The second structure (Figure 23b) is a thin-walled prismatic shell, hereafter referred to as the box, designed to resemble, in scale, the proportions of a washing machine. This choice allows for the study of a structure with different properties:

- similar to an industrially relevant radiating object;
- geometrically more complex and lacking an analytical solution.

The box is also made of steel, and it is built starting from three main components: a basement, a steel lateral surface, and a plastic top manufactured from *PLA* reinforced with carbon fibres. These three parts are clamped together, forming a closed thin-walled shell structure. In Figure 23b, the geometry of the lateral surface and the top are shown in an open configuration, with numerical labels assigned to the five faces for identification purposes, as well as the position of the excitation point. The main dimensions and properties of the box are provided in Table 3.

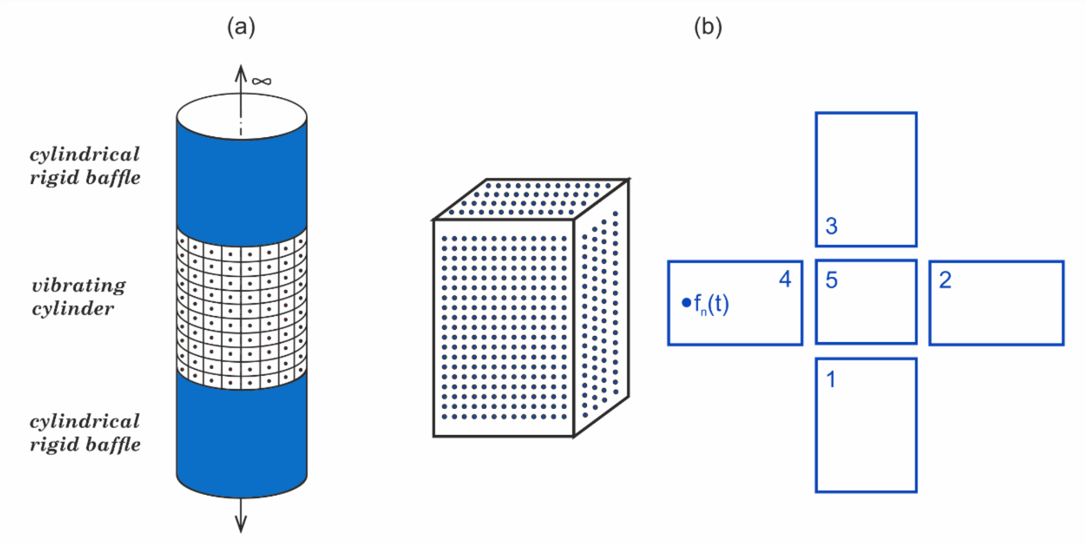


Figure 23: test structures considered: a) baffled cylinder and b) small box with ID numbers on each face

Table 3: thin-walled box general properties

BOX: General Properties		
Thickness	$2\ mm$	
Material lateral surfaces	Steel	
	Density	$\rho_b\ = 7500\ kg/m^3$
	Young modulus	$E_b\ = 21E^{10}\ N/m^2$
	Poisson ratio	$\nu_b\ = 0.29$
Material top surface	PLA	
	Density	$\rho_{PLA}\ = 1220\ kg/m^3$
	Young modulus	$E_{PLA}\ = 0.275E^{10}\ N/m^2$
	Poisson ratio	$\nu_{PLA}\ = 0.331$

Box face	**Dimension (mm)**	**N° target points**
Face 1	198x320	15x19
Face 2	240x320	12x19
Face 3	198x320	15x19
Face 4	240x320	12x19
Face 5	240x198	15x12

As seen for the cylinder, its faces are equipped with target points for *FRF* acquisition. These measurement points were subsequently used as nodes in a triangular mesh for surface discretisation. In this case too, the bending excitation was performed using a shaker fixed to the basement inside the box and connected to *face 4* via a *stinger*. Once again, the structure was equipped with an impedance head to record force measurements from the excitation and acceleration data.

Both geometries were mounted on a rotating platform, enabling rotation around their vertical axis to facilitate multi-view recordings using a single high-speed camera. The measurement grids on both structures were designed with a spatial resolution sufficient to capture the structural response across the investigated frequency range, generally in accordance with the criterion of at least six elements per half wavelength [88].

5.2 TEST RIGS: HIGH-SPEED CAMERA AND LASER VIBROMETER

Figure 24 illustrates the baffled cylindrical model structure and the box designed for this study. As shown in Figure 24b, the cylinder was mounted on two rigid flanges so that it can be considered constrained at both ends. Furthermore, as depicted in Figure 24a, two cylindrical extensions were attached to the top and bottom of the cylinder, each closed with circular end-caps. Figure 24d,e show the box structure placed directly on the floor; therefore, the sound it radiates will be affected by floor reflections. The bottom disk of the cylinder and the base of the box were mounted on a base flange using a rotational joint, enabling the entire structures to be rotated during the measurement campaigns while keeping both the optical camera and the laser vibrometer fixed, as mentioned previously. Figure 24b, c, d and f show the high-contrast speckle pattern applied to the cylinder and the box surfaces to facilitate the Digital Image Correlation image-processing stage. Two rows of ArUco markers were printed at the top and bottom of the pattern on the cylinder (Figure 24b) to enable the automatic extrinsic calibration of the multi-view imaging system. For the box (Figure 24d), calibration markers were arranged in two rows at the top and bottom of the lateral faces, while on the top face they were positioned along all four edges.

The flexural vibration reconstructed from the camera measurements were validated against reference measurements obtained using a laser vibrometer on both the test structures, as shown in Figures 25 and 27. To facilitate this validation process, the speckle pattern decal applied to the two different test structures, as illustrated in Figures 24c,f, was enhanced with a grid of green circular markers. These markers served as target points for the laser vibrometer measurements and they represent the centre of the pistons for the sound radiation derivation from the cylinder flexural vibration and the nodes of the triangular mesh for the sound radiation derivation from the box flexural vibrations.

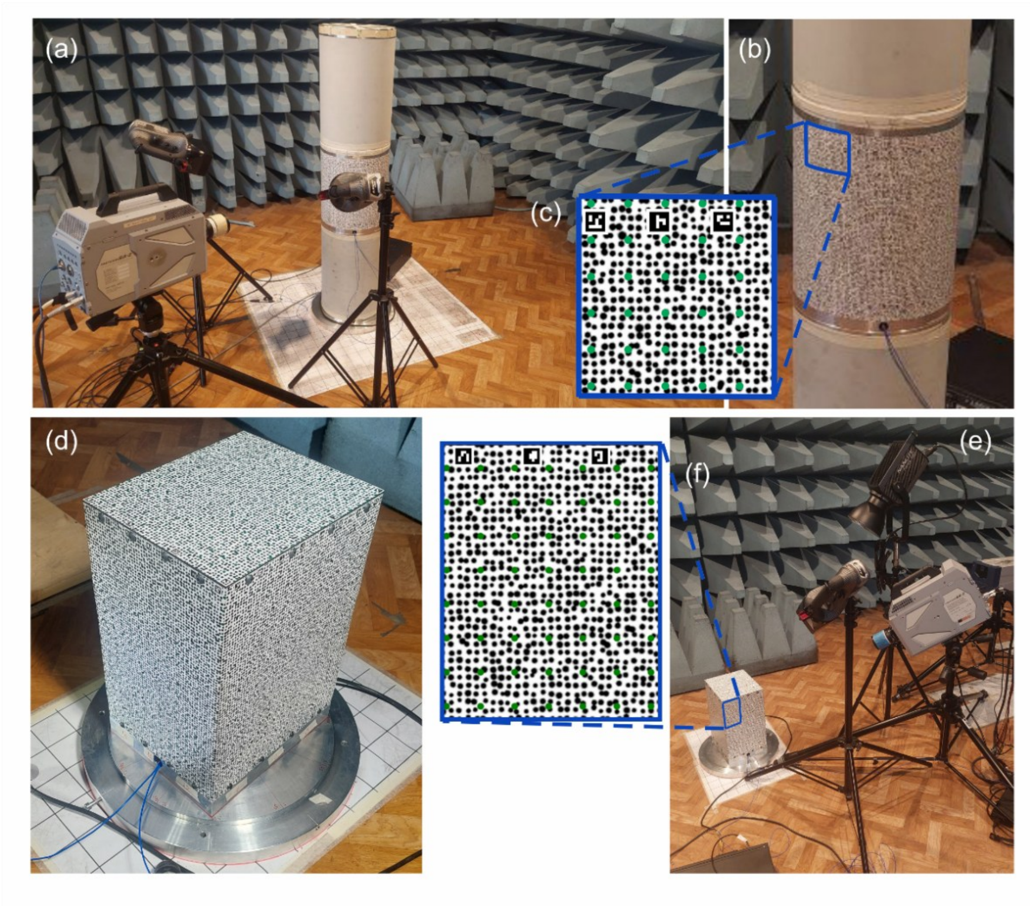


Figure 24: vibration measurements with high-speed camera test rigs: a) baffled cylinder in anechoic chamber with a single high-speed camera and two continuous light sources. b) vibrating cylinder with the speckle pattern for DIC purposes and c) magnification of the speckle pattern applied on the surface of the cylinder with ArUco markers and green dots for laser measurements. d) small thin-walled box structure with the speckle pattern applied. e) small box in a semi-anechoic chamber with a single high-speed camera and two continuous light sources and f) magnification of the speckle pattern applied on the surface of the box with ArUco markers and green dots for laser measurements.

MEASUREMENTS ON THE BAFFLED CYLINDER

The cylinder was rotated around its vertical axis between consecutive video acquisitions. In this way, high-speed camera footage of the cylinder under stationary broadband random excitation was acquired from 18 different viewpoints, using the same excitation signal in each measurement. The Photron FASTCAM SA-Z high-speed optical camera shown in Figure 24a was used to measure 50000 *12-bit* monochrome images (see an example in Figure 26a) at 10000 frames-per-second, with the spatial resolution of 896×896 pixels for each viewpoint, producing 1.31 terabytes of image-data in total.

Similarly, the flexural vibrations across the entire set of markers of the baffled cylinder structure, were acquired using the laser vibrometer as shown in Figure 25, which was oriented in such a way as its vertical measurement plane intersected the cylinder vertical axis. Specifically, the laser system automatically scanned the vibrations of each vertical array of 20 markers lying within its measurement plane. After each scan, the cylinder was incrementally rotated by steps of 6° to bring the adjacent vertical array of markers into alignment with the laser plane, repeating the process until the full surface was covered. The laser vibrometer recorded the mobility *FRF* spectra corresponding to the radial velocity responses per unit force excitation applied by the internal shaker.

Figure 25: laser vibrometer measurement setup

MEASUREMENTS ON THE BOX

The box was incrementally rotated around its vertical axis between successive video acquisitions. This procedure enabled the capture of high-speed camera footage of the box subjected to stationary broadband random excitation from 16 distinct viewing angles, with the same excitation signal consistently applied across all measurements. A Photron FASTCAM SA-Z high-speed optical camera shown in Figure 24e was employed to record 50000 monochrome images (*12-bit*) at a frame rate of 10000 frames per second, an example of recorded frame is reported in Figure 26b. Each recording had a spatial resolution of 896×896 pixels, resulting in a total dataset of approximately 1 terabyte of image data.

The flexural vibrations at all marker positions (reported in Table 3) were measured using the laser vibrometer as shown in Figure 27, which was oriented to capture the normal components of the flexural vibration on each face of the box.

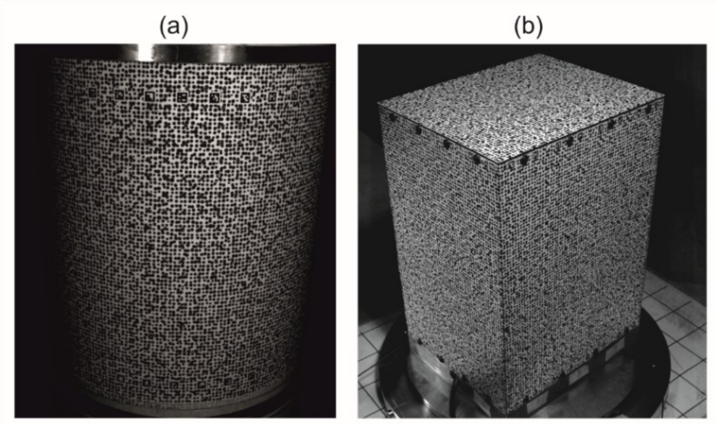


Figure 26: example of recorded frames: a) baffled cylinder and b) small box

The laser system performed an automated scanning of the vibration response for the matrix of markers on one of the five surfaces before the box was incrementally rotated to facilitate measurements on adjacent faces. Only for the measurement of the top surface was it necessary to reposition the laser to ensure that it remained perpendicular to the measurement surface. The laser vibrometer acquired the spectra of the mobility Frequency Response Functions for the normal velocities per unit force excitation, with the excitation provided by the shaker located inside the box.

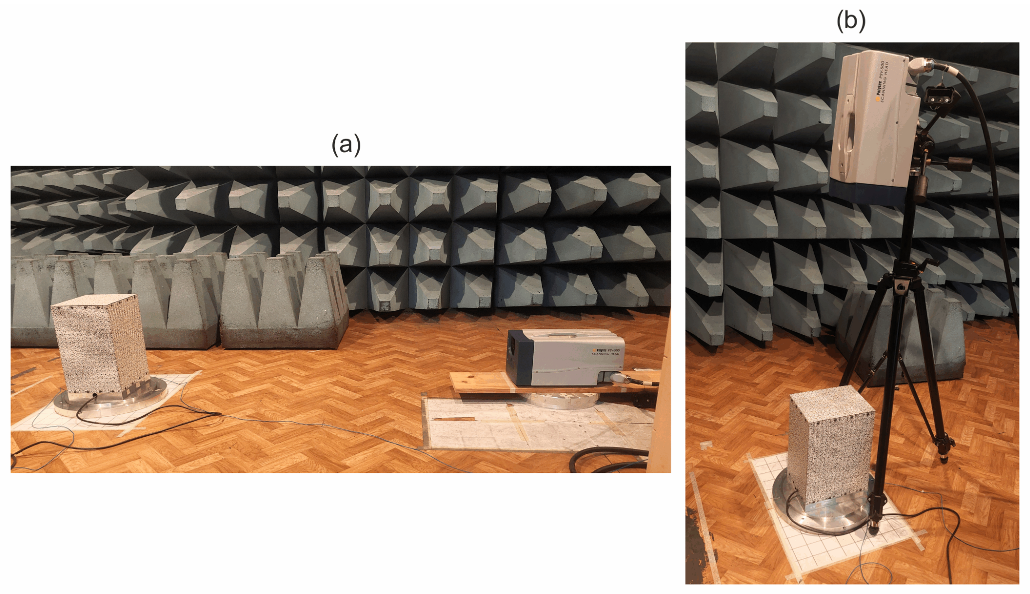


Figure 27: laser vibrometer measurement setup a) lateral faces and b) top face of the box

5.3 FREQUENCY DOMAIN TRIANGULATION

As previously outlined, this part of the study employs a frequency-domain triangulation approach to extract the flexural vibration field of a cylindrical and a prismatic structure from digital images acquired using a single camera. The adopted methodology, originally introduced in Ref. [39], enables the reconstruction of the entire vibrational response of the structure from image-based measurements without the need for multiple synchronized camera views, distinguishing it from conventional multi-camera techniques [63]. Unlike traditional multi-view triangulation, which requires simultaneous recordings from distinct camera positions [89], the present approach relies on a single camera capturing sequential images as the specimen is incrementally rotated between acquisitions. Throughout this process, the structure is excited by a stationary white noise force of sufficiently small amplitude to ensure a linear system response. The resulting image sequences are used to estimate displacements, which are then transformed into the frequency domain to yield spectral representations of the physical motion via frequency-domain triangulation [39]. The method assumes linear, time-invariant system behaviour and small-amplitude harmonic displacements, allowing for the reconstruction of spatial operating deflection shapes (*ODS*) using only a single moving camera.

As detailed in Chapter 3, the core workflow of the method is outlined as follows (also schematically represented in Figure 28):

a. Acquisition of image sequences with a rotating specimen subjected to stationary excitation.

b. Calibration of the multi-view imaging geometry, accounting for the rotation of the structure.

c. Estimation of $2D$ image-plane displacements via *Lucas–Kanade* digital image correlation [90], leading to the identification of displacement frequency response functions.

d. Application of frequency-domain triangulation to reconstruct $3D$ dynamic displacements and compute the associated dynamic flexibility *FRFs*.

These steps are necessary for the subsequent reconstruction of both the vibrational field and the associated acoustic radiation, which result from tonal excitations applied to the structure.

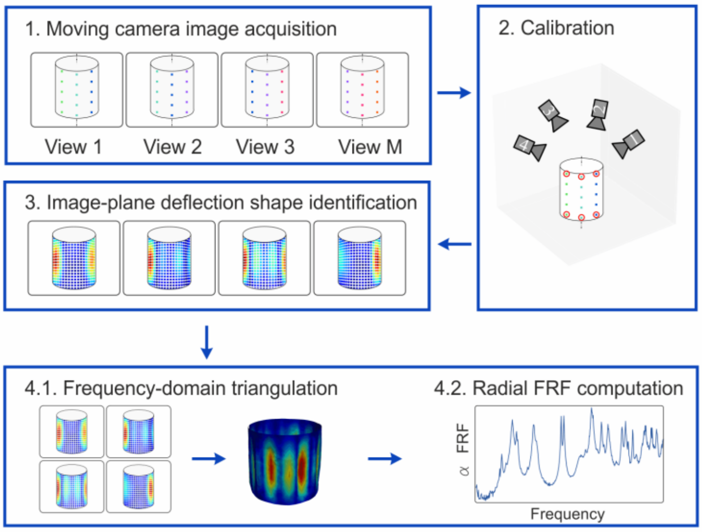


Figure 28: image-based 3D displacement measurement procedure using frequency-domain triangulation and the single-camera multi-view approach.

As introduced in prior chapters, the projection of a $3D$ point $\boldsymbol{M_i}$ (expressed in homogeneous coordinates) onto the image plane is described by the projection matrix $\boldsymbol{P}$ (Equation 2.13), yielding the corresponding $2D$ image point $\boldsymbol{m_i}$. In the present context, the objective is to use frequency-domain triangulation to reconstruct the dynamic flexibility *FRFs* between the radial displacements $u_r^n(\omega)$ of the cylinder, measured at a set of marker locations n, and the radial force input $f_r(\omega)$ applied to the structure.

Assuming time-harmonic excitation, the radial displacement and force are expressed in complex form as:

$$\begin{cases} u_r(\boldsymbol{m}_n, t) = \mathcal{R}e\{u_r(\boldsymbol{m}_n, \omega)e^{i\omega t}\} \\ f_r(\boldsymbol{m}_f, t) = \mathcal{R}e\{f_r(\boldsymbol{m}_f, \omega)e^{i\omega t}\} \end{cases} \tag{5.1}$$

where ω denotes the angular frequency, $i = \sqrt{-1}$, $u_r(\boldsymbol{m}_n, \omega)$ represents the complex amplitude of the radial displacement at the $n-th$ marker position, and $f_r(\boldsymbol{m}_f, \omega)$ is the complex amplitude of the applied radial force at position $\boldsymbol{m}_f$.

The dynamic flexibility *FRF* between the input force and the displacement responses at the marker locations is then defined by:

$$\alpha_{r,n}(\omega) = \frac{u_r(\boldsymbol{m}_n, \omega)}{f_r(\boldsymbol{m}_f, \omega)}. \tag{5.2}$$

The goal of the triangulation procedure is to retrieve the complex amplitudes of the radial displacements at the centres of the quadrilateral regions used as elemental acoustic radiators. To compute the receptance *FRF* defined in Equation (5.2), the radial component of each displacement vector $u_r(\boldsymbol{m}_n, \omega)$ is extracted through a projection onto the unit surface normal $\boldsymbol{n}_n$ at the marker location $\boldsymbol{m}_n$, using:

$$u_r(\boldsymbol{m}_n, \omega) = \frac{u_r(\boldsymbol{m}_n, \omega)\boldsymbol{n}_n}{|\boldsymbol{n}_n|}. \tag{5.3}$$

Analogously, the complex amplitude of the force excitation is computed from its time-domain signal by applying a discrete Fourier transform (*DFT*) over a finite set of samples:

$$f_r(\boldsymbol{m}_f, \omega) = \frac{1}{N}\sum_{n=1}^{N} f_r(\boldsymbol{m}_f, t_n)e^{-2i\pi nk/N}. \tag{5.4}$$

5.4 STRUCTURAL NATURAL MODES OF CYLINDRICAL SHELLS

In the free vibration case, the equation of motion of a cylindrical shell structure simply supported can be written as [58,91–94]:

$$\begin{gathered} D\left[\frac{\partial^8 u_r}{\partial z^8} + 4\frac{\partial^8 u_r}{\partial z^6 R^2 \partial\theta^2} + 6\frac{\partial^8 u_r}{\partial z^4 R^4 \partial\theta^4} 4\frac{\partial^8 u_r}{\partial z^2 R^6 \partial\theta^6} + \frac{\partial^8 u_r}{R^8 \partial\theta^8}\right] \\ + \frac{E_c s_c}{R^2}\frac{\partial^4 u_r}{\partial z^4} + \rho_c s_c\left[\frac{\partial^4 \ddot{u}_r}{\partial z^4} + 2\frac{\partial^4 \ddot{u}_r}{\partial z^2 R^2 \partial\theta^2} + \frac{\partial^4 u_r}{R^4 \partial\theta^4}\right] = 0 \end{gathered} \tag{5.5}$$

Where $D = \frac{E_c s_c}{1-\nu^2}$ represents the bending stiffness of the shell and $\ddot{u}_r$ is the second derivative of the radial displacements and z, R, θ represents the coordinate system as shown in Figure29.

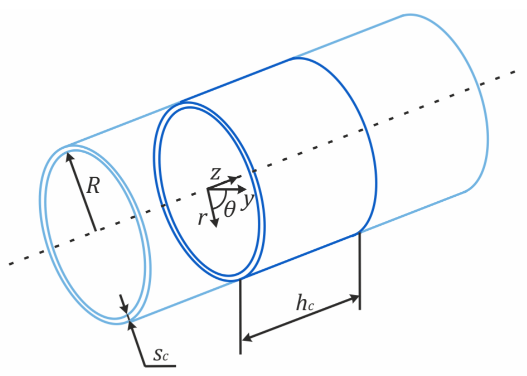


Figure 29: sketch of a thin-walled cylinder with nomenclature

For a simply supported cylinder the following constraints are set to the boundaries:

$$\begin{cases} u_r(0,\theta,t) = 0 \\ u_r(h_c,\theta,t) = 0 \\ M_z(0,\theta,t) = \left[-\dfrac{Es_c^3}{12(1-\nu^2)}\left(-\dfrac{\partial^2 u_r}{\partial z^2} - \nu\dfrac{\partial^2}{R^2\partial\theta^2}\right)\right]_{z=0} = 0 \\ M_z(h_c,\theta,t) = \left[-\dfrac{Es_c^3}{12(1-\nu^2)}\left(-\dfrac{\partial^2 u_r}{\partial z^2} - \nu\dfrac{\partial^2}{R^2\partial\theta^2}\right)\right]_{z=h_c} = 0 \end{cases} \tag{5.6}$$

and after some mathematical manipulations the Equation (5.5) can be simplified and rewritten as:

$$D\left[k_z^2 + \frac{k_\theta^2}{R^2}\right]^4 + \frac{E_c s_c}{R^2}k_z^4 - \rho_c s_c\left[k_z^2 + \frac{k_\theta^2}{R^2}\right]^2 \omega_c^2 = 0, \tag{5.7}$$

where k_z and k_θ can be obtained by considering the general solution to equations [91]:

$$\begin{cases} k_z = -\dfrac{1}{\mathcal{X}(z)}\dfrac{d^2\mathcal{X}(z)}{dz^2} = \dfrac{m_1\pi}{h_c} \\ k_\theta = -\dfrac{1}{\mathcal{Q}(\theta)}\dfrac{d^2\mathcal{Q}(\theta)}{d\theta^2} = m_2 \end{cases} \tag{5.8}$$

which are of the form:

$$\begin{cases} \mathcal{X}(z) = A_1\sin(k_z z) \\ \mathcal{Q}(\theta) = B_1\cos(k_\theta\theta) + B_2\sin(\mathrm{k_\theta\theta}) \end{cases} \tag{5.9}$$

where A_1, B_1 and B_2 are arbitrary constants.

From Equation (5.7) it is possible to obtain the angular frequency of the free vibration in terms of the axial and circumferential wavenumbers k_z and k_θ:

$$\omega_c^2 = \frac{E_c}{\rho R^2}\left\{\frac{(Rk_z)^4}{[(Rk_z)^2 + k_\theta^2]^2} + \frac{s_c^2/R^2}{12(1-\nu^2)}\left[(Rk_z)^2 + k_\theta^2\right]^2\right\} \tag{5.10}$$

and substituting Equation (5.8) in Equation (5.10) the angular frequency in terms of the modal indices m_1 and m_2 is obtained [91,94]:

$$\omega_c^2 = \frac{E_c}{\rho R^2}\left\{\frac{\left(\frac{R m_1 \pi}{h_c}\right)^4}{[\left(\frac{R m_1 \pi}{h_c}\right)^2 + m_2^2]^2} + \frac{s_c^2/R^2}{12(1-\nu^2)}\left[\left(\frac{R m_1 \pi}{h_c}\right)^2 + m_2^2\right]^2\right\}, \tag{5.11}$$

at which two distinct flexural natural modes occur, namely the symmetric $\varphi^s_{m_1,m_2}(z)$ and anti-symmetric $\varphi^a_{m_1,m_2}(z)$ modes:

$$\begin{cases} \varphi^s_{m_1,m_2}(z) = \sin\left(\frac{m_1\pi}{h_c}z\right)\cos(m_2\theta) \\ \varphi^a_{m_1,m_2}(z) = \sin\left(\frac{m_1\pi}{h_c}z\right)\sin(m_2\theta) \end{cases} \tag{5.12}$$

A commonly adopted representation of the natural frequencies of a cylindrical shell structure [91,94] is shown in Figure 30, where the natural frequency, derived from Equation (5.11), is plotted as a function of the circumferential modal index m_2, for different values of the axial index m_1. This diagram reveals a key feature of simply supported cylindrical shells: the lowest natural frequency does not necessarily correspond to the lowest values of the modal indices, as is typically the case for flat rectangular plates. Specifically, when the circumferential index $m_2 = 0$, the associated flexural displacement is independent of the angular coordinate.

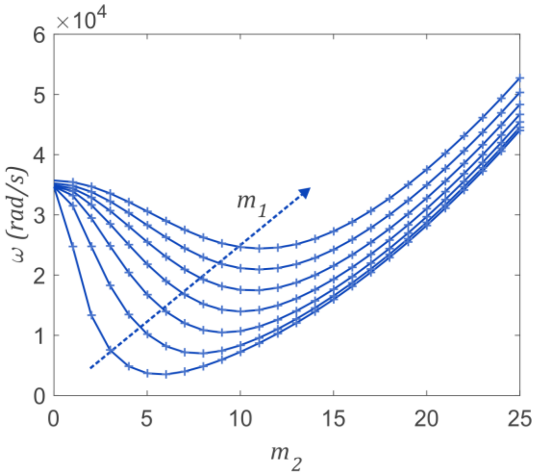


Figure 30: natural frequency of a simple supported cylinder as a function of the circumferential modal index m_2

In terms of nodal patterns, the following relationships hold:

- The number of axial nodal lines is equal to $m_1 - 1$;

- The number of circumferential nodal lines is $2m_2$.

Alternatively, the modal indices can also be interpreted in terms of wavelength content:

- The axial modal index m_1 corresponds to the number of axial half-wavelengths;

- The circumferential modal index m_2 indicates the number of circumferential full wavelengths.

5.5 STRUCTURAL NATURAL MODES OF THIN PLATES

A box structure can be approximated as multiple plates connected together. Thus, thin plates theory is here presented. A thin plate can be considered as a shell with zero curvatures. In the free vibration case, the equation of motion can be written as [91–93,95,96]:

$$D\left(\frac{\partial^4 u_n}{\partial x^4} + 2\frac{\partial^4 u_n}{\partial x^2 \partial y^2} + \frac{\partial^4 u_n}{\partial y^4}\right) + \rho_b s_b \ddot{u}_n = 0, \tag{5.13}$$

where $D = \frac{E_b s_b^3}{12(1-\nu^2)}$ represents the bending stiffness of the shell and $\ddot{u}_n$ is the second derivative of the radial displacements. For the plate with $l \times h$ shown in Figure 31, with two opposite edges simply supported (SS) the following constraints are set to the boundaries:

$$\begin{cases} u_n(0,y,t) = 0 \\ u_n(h,y,t) = 0 \\ M_{xx}(0,y,t) = \left[-EI\dfrac{\partial^2 u_n}{\partial x^2}\right]_{x=0} = 0 \\ M_{xx}(h,y,t) = \left[-EI\dfrac{\partial^2 u_n}{\partial x^2}\right]_{x=h} = 0 \end{cases} \tag{5.14}$$

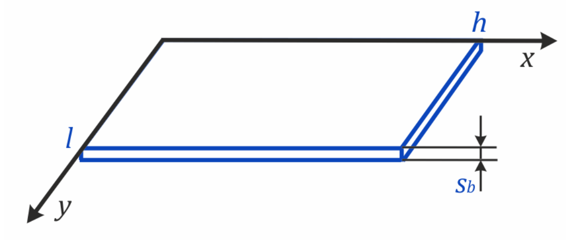


Figure 31: sketch of a thin plate with nomenclature

Equation (5.13) can be rewritten considering complex notation in [91,96]:

$$D\left(\frac{\partial^4 U_n}{\partial x^4} + 2\frac{\partial^4 U_n}{\partial x^2 \partial y^2} + \frac{\partial^4 U_n}{\partial y^4}\right) - \omega_p^2 \rho_b s_b U_n = 0, \tag{5.15}$$

where $u_n(x, y, t) = U_n(x, y)e^{i\omega t}$. And after some mathematical manipulations the Equation (5.13) can be simplified and rewritten in:

$$\frac{d^4\mathcal{Y}}{dy^4} - 2\left(\frac{m\pi}{h}\right)^2 \frac{d^2\mathcal{Y}}{dy^2} + \left[\left(\frac{m\pi}{h}\right)^4 - \frac{\rho_b s_b}{D}\omega_p^2\right]\mathcal{Y} = 0. \tag{5.16}$$

Here $\mathcal{Y} = U_n/\sin\left(\frac{m\pi x}{h}\right)$. Therefore, applying the boundary conditions in (5.14), the solution to Equation (5.16) can be written as:

$$\left(\frac{\lambda_i}{l}\right)^4 - 2\left(\frac{m\pi}{h}\right)^2\left(\frac{\lambda_i}{l}\right)^2 + \left[\left(\frac{m\pi}{h}\right)^4 - \frac{\rho_b s_b}{D}\omega_p^2\right] = 0, \tag{5.17}$$

which solved for λ_i gives:

$$\lambda_i = \pm\frac{l}{h}m\pi\sqrt{1 \pm \frac{\omega_p}{(m^2\pi^2/h)\sqrt{D/\rho_b s_b}}} = \pm\frac{l}{h}m\pi\sqrt{1 \pm K}, \tag{5.18}$$

in which $K > 1$. After some mathematical manipulation the procedure leads to:

$$\begin{aligned}\mathcal{Y}(y) = A_1 \cosh\left(\lambda_1 \frac{y}{l}\right) + A_2 \sinh\left(\lambda_1 \frac{y}{l}\right) \\ + \frac{A_3}{i}\cos\left(\lambda_3 \frac{y}{l}\right) + A_4 \sin\left(\lambda_3 \frac{y}{l}\right).\end{aligned} \tag{5.19}$$

Considering now the other to edges clamped (CC), it is possible to add the following two constraints at the boundaries [91,92]:

$$\begin{cases} u_n(x,0,t) = 0 \\ u_n(x,l,t) = 0 \\ \dfrac{\partial u_n}{\partial y}(x,0,t) = 0 \\ \dfrac{\partial u_n}{\partial y}(x,l,t) = 0 \end{cases} \tag{5.20}$$

and solving for A_1, A_2, A_3, A_4 and obtaining the expression for $K_n(n = 1,2,3, \dots)$, the frequency of a plate simply supported at two opposite edges and clamped at the other two can be written as:

$$\omega_p = \frac{m^2\pi^2 K_n}{h^2\sqrt{\frac{\rho_b s_b}{D}}} \tag{5.21}$$

and the mode shape is obtained by letting:

$$U_n(x,y) = \mathcal{Y}(y)\sin\left(\frac{m\pi x}{h}\right). \tag{5.22}$$

The natural frequencies of a $CSCS$ plate derived from Equation (5.21) are shown in Figure 32, where the natural frequency is plotted as a function of the natural frequency dimensionless parameter $\lambda_{CSCS}(boundary\ condition) = \lambda_{mn}$,[92] for different values of h.

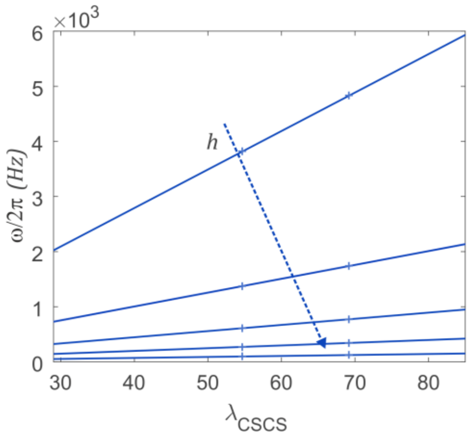


Figure 32: CSCS plate natural frequency as a function of dimensionless parameter λ_{CSCS}

To better illustrate the influence of m and n, the derivation for an $SSSS$ plate is provided. Therefore, if the boundary conditions in (5.20) are changed to simply supported (SS), the expressions for natural frequency and mode shapes become more easily derivable and readable, and for an $SSSS$ plate they are [91,92,95,96]:

$$\omega_p = \pi^2 \left[\left(\frac{m}{h}\right)^2 + \left(\frac{n}{l}\right)^2 \right] \sqrt{\frac{D}{\rho_b s_b}} \tag{5.23}$$

$$U_n(x,y) = Dsin\left(\frac{n\pi y}{l}\right) \sin\left(\frac{m\pi x}{h}\right). \tag{5.24}$$

The natural frequencies of a $SSSS$ plate derived from Equation (5.23) are shown in Figure 33, where the natural frequency is plotted as a function of the modal index n, for different values of the index m.

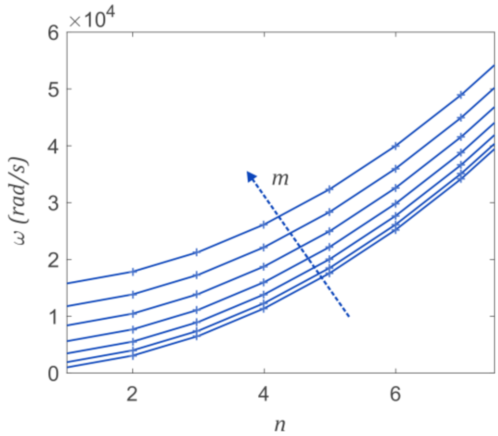


Figure 33: SSSS plate natural frequency as a function of the modal index n

5.6 VIBRATION FIELD AND KINETIC ENERGY SPECTRA

To start with, the vibration response of the two model structures is considered here. The total flexural response of structures can be conveniently derived in terms of the total flexural kinetic energy. For stationary time-harmonic processes, the total flexural vibration of the cylinder and the box structures can be conveniently derived in terms of the time averaged total flexural kinetic energy:

$$\bar{K} = \lim_{T\to\infty} \frac{1}{\mathrm{T}} \int_0^T K(t)\mathrm{d}t, \tag{5.25}$$

where, for a thin-walled structure (as a cylinder or a plate), the instantaneous total flexural kinetic energy is given by:

$$K(t) = \frac{1}{2}\rho s \int_{S_c} \dot{u}(\boldsymbol{x}, t)^2 \mathrm{dS}. \tag{5.26}$$

Thus, for time-harmonic velocities characterised by complex amplitude $\dot{u}(\boldsymbol{x}, \omega) = i\omega u(\boldsymbol{x}, \omega)$, it is obtained:

$$\overline{K}(\omega) = \frac{1}{4}\rho s \int_{S_c} |\dot{u}(\boldsymbol{x}, \omega)|^2 \mathrm{d}S. \tag{5.27}$$

Here, ρ is the density of the material (the cylinder or the box) having thickness s and surface area S.

CYLINDER FLEXURAL KINETIC ENERGY SPECTRA

Starting from the cylinder, the integral expression in Equation (5.27) can be approximated by a Riemann sum over the regular mesh of $N_c = 62 \times 20 = 1240$ quadrilateral elements into which the surface of the cylinder was divided to, such that:

$$\overline{K}_c(\omega) \cong \frac{\rho_c s_c S_c}{4N_c} \sum_{\mathrm{n}=1}^{N_c} |\dot{u}_{rn}(\omega)|^2, \tag{5.28}$$

where $\dot{u}_{rj}(\omega)$ is the complex amplitude of the radial velocity at the centre of the *j-th* element and S_c/N_c is the area of the quadrilateral elements which are identical.

At this point, recalling the formulation presented above for the reconstruction from multi-view camera measurements of the complex amplitudes of the radial vibration displacements at the centres of the mesh of 62×20 quadrilateral elements, the time-averaged kinetic energy at the circular frequency ω_k can be expressed in terms of the dynamic flexibility defined in Equation (5.2), such that:

$$\overline{K}_c(\omega_k) \cong \omega^2 \frac{\rho_c s_c \mathrm{S_c}}{4N_c} \sum_{\mathrm{n}=1}^{N_c} |\alpha_{rn}(\omega_k)|^2 |f_r(\omega_k)|^2, \tag{5.29}$$

where $|f_r(\omega_k)|$ is the amplitude of the radial force exerted by the shaker, which was kept to a constant value f_r in the frequency range of interest. In practice, this summation is implemented with a vector formulation. To this end, the complex displacements and the dynamic flexibility *FRFs* are casted into the following column vectors $\mathbf{u_r}(\omega_k) = \lfloor u_{r_1}(\omega_k) \quad \cdots \quad u_{rN}(\omega_k) \rfloor^{\mathrm{T}}$ and $\boldsymbol{\alpha}_\mathbf{r}(\omega_k) = \lfloor \alpha_{\mathrm{r}1}(\omega_k) \quad \cdots \quad \alpha_{rN}(\omega_k) \rfloor^{\mathrm{T}}$, such that:

$$\mathbf{u_r}(\omega_k) = \boldsymbol{\alpha}_\mathbf{r}(\omega_k) f_r. \tag{5.30}$$

Hence, recalling Equation (5.29), the flexural kinetic energy for the cylinder can be written as:

$$
\begin{aligned}
\overline{K}_c(\omega_k) &\cong \omega^2 \frac{\rho_c s_c \mathrm{S_c}}{4N_c} \mathbf{u_r}(\omega_k)^{\mathrm{H}} \mathbf{u_r}(\omega_k) \\
&= \omega^2 \frac{\rho_c s_c \mathrm{S_c}}{4N_c} \boldsymbol{\alpha}_{\mathbf{r}}(\omega_k)^{\mathrm{H}} \boldsymbol{\alpha}_{\mathbf{r}}(\omega_k) |f_r(\omega_k)|^2.
\end{aligned} \tag{5.31}
$$

As discussed in Section 4.2 and shown in Figure 25, the validity of the vibration measurement approach based on multi-view video recordings proposed in this thesis has been assessed with measurements taken with a scanning laser vibrometer measurements. As depicted in Figure 25, the laser vibrometer has been used to measure mobility *FRFs* given by the ratios between the complex amplitudes of the velocities measured at the grid of centre points, $\dot{u}_{rn}(\omega_k)$, and the complex amplitude of the radial force exerted on the cylinder, $f_r(\omega_k)$:

$$
Y_{rj}(\omega_k) = \frac{\dot{u}_{rj}(\omega_k)}{f_r(\omega_k)}. \tag{5.32}
$$

Hence, recalling Equation (5.28), the time averaged kinetic energy has been derived with the following formula:

$$
\overline{K}_c(\omega_k) \cong \frac{\rho_c s_c \mathrm{S_c}}{4N_c} \sum_{j=1}^{N_c} \left| Y_{rj}(\omega_k) \right|^2 f_r^2. \tag{5.33}
$$

Here too, the complex velocity and the mobility FRFs have been casted into the following column vectors $\dot{\mathbf{u}}_r(\omega_k) = \lfloor \dot{u}_{\mathrm{r}1}(\omega_k) \quad \cdots \quad \dot{u}_{rN}(\omega_k) \rfloor^{\mathrm{T}}$ and $\mathbf{Y_r}(\omega_k) = \lfloor Y_{\mathrm{r}1}(\omega_k) \quad \cdots \quad Y_{rN}(\omega_k) \rfloor^{\mathrm{T}}$, such that:

$$
\dot{\mathbf{u}}_r(\omega_k) = \mathbf{Y_r}(\omega_k) f_r \tag{5.34}
$$

and thus:

$$
\overline{K}_c(\omega_k) \cong \frac{\rho_c s_c \mathrm{S_c}}{4N_c} \dot{\boldsymbol{u}}_r(\omega_k)^H \dot{\boldsymbol{u}}_r(\omega_k) = \frac{\rho_c s_c \mathrm{S_c}}{4N_c} \mathbf{Y}_{\mathbf{r}}^{\mathrm{H}}(\omega_k) \mathbf{Y_r}(\omega_k) f_r^2. \tag{5.35}
$$

Figure 34 shows the spectrum of the time-averaged total flexural kinetic energy per unit force excitation derived with Equation (5.35) from the laser vibrometer measurements in a frequency range comprised between 500 and $1500\ Hz$. The figure encompasses small pictures with the deflection shapes at the resonance frequencies. In general, the spectrum is characterised by distinct resonance peaks, which suggests that in the $500\ -$

$1500\ Hz$ range the flexural response is characterised by a low modal overlap [97]. Indeed, apart from few cases, the deflection shapes at the resonance frequencies show the typical regular pattern for the flexural natural modes of a cylinder [91,94,98,99].

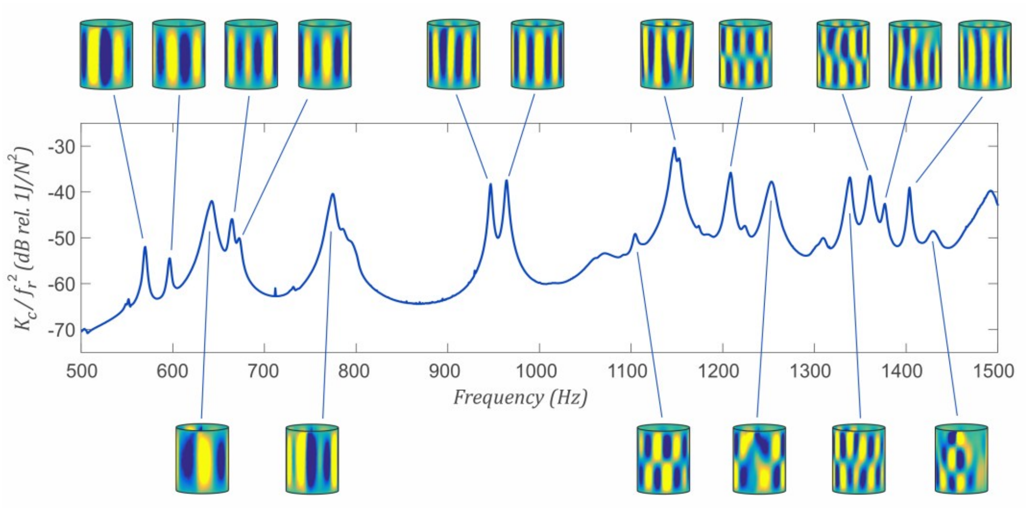


Figure 34: spectrum of the total flexural kinetic energy per unit force excitation derived from laser vibrometer measurements with the deflection shapes at the resonance frequencies.

In general, the deflection shapes show quite high circumferential mode orders that span from 5 to 11 and rather low axial mode orders equal to 1 or 2. In the latter case, as can be noticed in Figure 34, the mode shapes are characterised by a circumferential nodal line exactly in the middle span of the cylinder height. There is just a deflection shape at about $1440\ Hz$, which shows a mode order 3 in the axial direction.

Moving to the flexural vibrations reconstructed from the camera measurements, the orange line in Figure 35 shows the spectrum of the time-averaged total flexural kinetic energy per unit force excitation derived with the formulation presented above and then calculated with Equation (5.31) in the $500 - 1500\ Hz$ frequency range. To allow a direct comparison, the spectrum obtained from the laser vibrometer measurements is reported as well with the blue line. The two spectra overlap quite well. There are small shifts in a few resonance frequencies and also small mismatches in the amplitudes of the resonance peaks. These small gaps are most likely due to the fact that the measurements had not been taken simultaneously. Consequently, the two measurement techniques are subjected to different environmental conditions, particularly in terms of temperature variations. It should be noted that the measurement performed with the laser vibrometer took several hours (for the cylinder, for example, around 4 hours, depending on the number of measurement points and the settings used in the FRF calculation software) such that the response of the cylinder was affected by a significant variability induced by ambient temperature changes and heating effects produced by the shaker located inside the cylinder. In contrast, measurements carried out using a single high-speed camera required only approximately 10 seconds of acquisition per view and were therefore less affected by changes in environmental boundary conditions over time. However, the experimental setup adopted for the camera measurement included two continuous light sources to ensure adequate illumination and sufficient contrast of the speckle pattern applied to the structure. The use of these light sources may have influenced the average surface temperature of the measured object. The spectrum obtained from the camera measurements is slightly noisier that that derived from the laser vibrometer measurements, particularly at the higher frequency end of the measurement range.

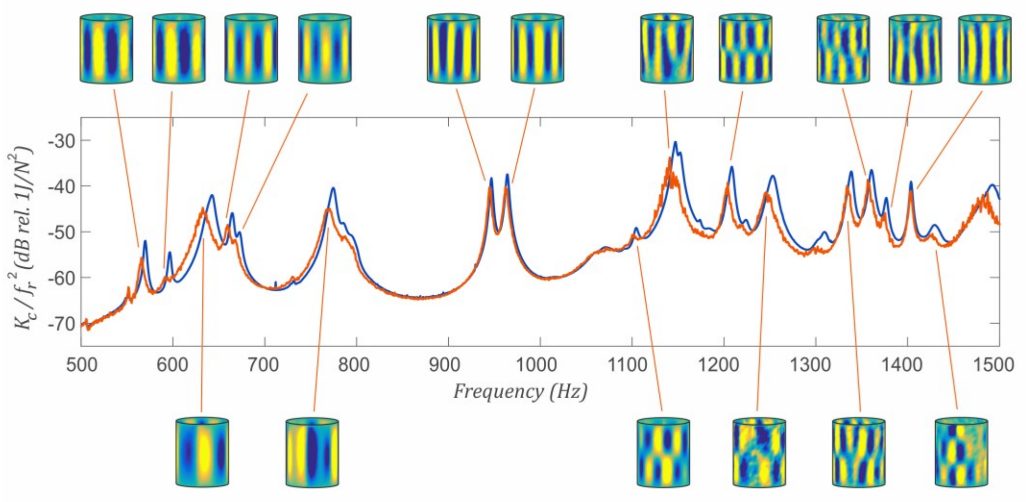


Figure 35: spectrum of the total flexural kinetic energy per unit force excitation derived from camera measurements (orange line) with the deflection shapes at the resonance frequencies (compared with laser vibrometer measurements, blue line).

Nevertheless, the two measurements show quite a remarkable global agreement, which is confirmed also by the deflection shapes reported in the two figures for the vibrometer and camera measurements respectively. The deflection shapes at higher resonance frequencies depicted in Figure 35 show a comparatively more irregular shape than those plotted in Figure 34. This indicates that indeed, the results obtained from the camera measurements are slightly less accurate than those obtained from the laser vibration measurements, although as will be shown in Chapter 6, and can already be seen in the spectral analysis, they have little impact on the overall vibration and thus on the sound radiation.

BOX FLEXURAL KINETIC ENERGY SPECTRA

Moving on to the box, the integral expression for the total flexural kinetic energy given in Equation (5.27) can be approximated by a Riemann sum over the regular mesh of N_n quadrilateral elements (reported in Table 3) where n indicates which face of the box is considered, such that:

$$\overline{\boldsymbol{K}}_{\boldsymbol{nb}}(\omega) \cong \frac{\rho_b s_b S_{nb}}{4N_n} \sum_{\mathrm{n}=1}^{N_n} |\dot{u}_{rn}(\omega)|^2. \tag{5.36}$$

Considering the box structure composed as 5 plates connected together, the total kinetic energy of the box was then approximated by:

$$\begin{aligned} \overline{\boldsymbol{K}}_{\boldsymbol{b}}(\omega) &\cong \sum_{n=1}^{5} \frac{\rho_b s_b S_{nb}}{4N_n} \sum_{\mathrm{n}=1}^{N_n} |\dot{u}_{rn}(\omega)|^2 \\ &= \overline{\boldsymbol{K}}_{\boldsymbol{1b}}(\omega) + \overline{\boldsymbol{K}}_{\boldsymbol{2b}}(\omega) + \overline{\boldsymbol{K}}_{\boldsymbol{3b}}(\omega) + \overline{\boldsymbol{K}}_{\boldsymbol{4b}}(\omega) + \overline{\boldsymbol{K}}_{\boldsymbol{5b}}(\omega) \end{aligned} \tag{5.37}$$

Where $\overline{\boldsymbol{K}}_{\boldsymbol{1b}}(\omega)$, $\overline{\boldsymbol{K}}_{\boldsymbol{2b}}(\omega)$, $\overline{\boldsymbol{K}}_{\boldsymbol{3b}}(\omega)$, $\overline{\boldsymbol{K}}_{\boldsymbol{4b}}(\omega)$ and $\overline{\boldsymbol{K}}_{\boldsymbol{5b}}(\omega)$ denote respectively the total flexural kinetic energy associated with each of the five distinct faces of the box which is shown in Figure 36 using blue colour for laser measurements and orange colour for camera measurements. Figure 36 shows also the main characteristics of each face of the box, in fact, it is possible to notice that the amplitudes of the resonance peaks are in a similar range for each face. Furthermore, it is possible to notice that generally, laser and camera measurements agree quite well. It is also observed that, at higher frequencies, the progressively reduced displacement amplitudes of the structure lead the camera-based measurements to exhibit greater difficulty in accurately reconstructing the spectra. The last spectrum reported in Figure 36 represents the total flexural kinetic energy of the top face of the box, which is recalled being manufactured using a different material with respect to the other faces. For this face, a more marked deviation between the curve obtained from laser vibrometer measurements and that obtained from camera measurements can be observed starting from a certain frequency range. This effect may be related to the material properties and to the construction and connection of this face to the remaining structure, potentially leading to a higher equivalent stiffness. This leads to even smaller displacement amplitudes compared to those observed on the other faces, resulting in a less accurate reconstruction when using the camera-based measurement technique.

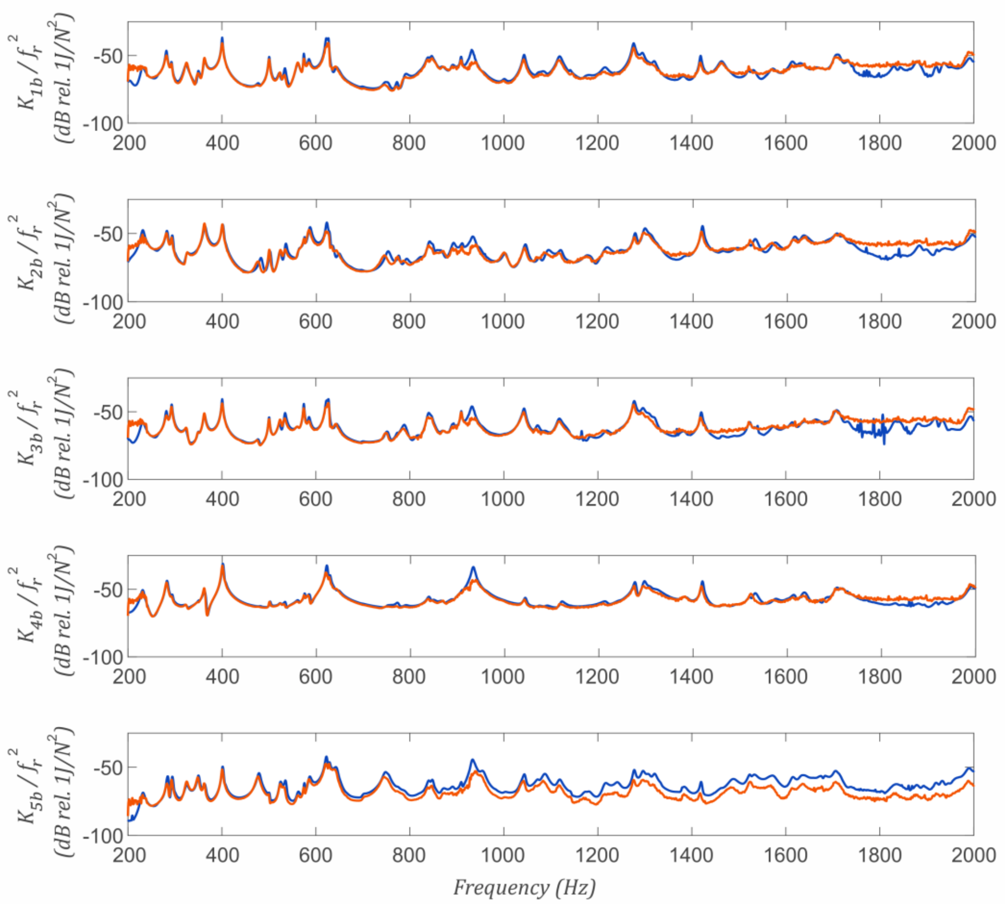


Figure 36: total flexural kinetic energy of each face of the box structure. Blue line laser vibrometer data and orange line high-speed camera data.

At this point, recalling the formulation presented above for the reconstruction from multi-view camera measurements of the complex amplitudes of the radial vibration displacements at the nodes of the mesh of elements, the time-averaged kinetic energy at the circular frequency ω_k can be expressed in terms of the dynamic flexibility defined in Equation (5.2), such that:

$$\overline{\boldsymbol{K}}_{\boldsymbol{b}}(\omega_k) \cong \sum_{n=1}^{5} \omega^2 \frac{\rho_b s_b S_b}{4N_n} \sum_{\mathrm{n}=1}^{N_n} |\alpha_{rn}(\omega_k)|^2 |f_r(\omega_k)|^2, \qquad (5.38)$$

where $|f_r(\omega_k)|$ is the amplitude of the radial force exerted by the shaker, which was kept to a constant value f_r in the frequency range of interest. Now using the same procedure adopted for the flexural kinetic energy derivation of the cylinder, the expression of Equation (5.38) can be written as:

$$\bar{K}_b(\omega_k) \cong \sum_{n=1}^{5} \omega^2 \frac{\rho_b s_b S_b}{4N_n} \mathbf{u}_\mathbf{r}(\omega_k)^\mathrm{H} \mathbf{u}_\mathbf{r}(\omega_k)$$
$$= \sum_{n=1}^{5} \omega^2 \frac{\rho_b s_b S_b}{4N_n} \boldsymbol{\alpha}_\mathbf{r}(\omega_k)^\mathrm{H} \boldsymbol{\alpha}_\mathbf{r}(\omega_k) |f_r(\omega_k)|^2. \tag{5.39}$$

Once more, as discussed in Section 5.2 and shown in Figure 27, the validity of the vibration measurement approach based on multi-view video recordings proposed in this thesis has been assessed with measurements taken with a scanning laser vibrometer. As depicted in Figure 27, and already presented for the cylinder, the laser vibrometer has been used to measure mobility *FRFs* $Y_{rj}(\omega_k)$ given by the ratios between the complex amplitudes of the velocities measured at the grid of nodal points, $\dot{u}_{rn}(\omega_k)$, and the complex amplitude of the radial force exerted on the box, $f_r(\omega_k)$. With the same assumptions made for the cylinder, the following formulation for the flexural kinetic energy of the box is obtained:

$$\bar{K}_b(\omega_k) \cong \sum_{n=1}^{5} \frac{\rho_b s_b S_b}{4N_n} \dot{\boldsymbol{u}}_{\boldsymbol{r}}(\omega_k)^H \dot{\boldsymbol{u}}_{\boldsymbol{r}}(\omega_k)$$
$$= \sum_{n=1}^{5} \frac{\rho_b s_b S_b}{4N_n} \mathbf{Y}_\mathbf{r}^\mathrm{H}(\omega_k) \mathbf{Y}_\mathbf{r}(\omega_k) f_r^2. \tag{5.40}$$

Figure 37 displays the spectrum of the time-averaged total flexural kinetic energy per unit force excitation for the box structure, obtained from the laser vibrometer measurements and calculated using Equation (5.40) over the frequency range from 200 to 2000 Hz. As in the case of the cylindrical shell discussed above, the spectrum reveals several well-defined resonance peaks, indicating that, also for the box, the flexural dynamic response in this frequency band is characterised by a relatively low modal overlap. In this case, the resonance peaks are denser in the frequency range comprised between 200 and 600 Hz. The figure also includes small snapshots of the deflection shapes at the resonance frequencies, which provide a qualitative insight into the modal behaviour of the structure. In general, the deflection shapes show low horizontal mode orders for the single plates, from 1 to 4 and low vertical mode orders equal to 1, 2, 3 or 4. However, for the small dimensions of the box (see Table 3), analysing frequencies above 1000 Hz, the mode shapes become progressively more complex, characterized by the presence of multiple lobes concentrated in a small area. As can be noticed in Figure 37, the mode shapes are described by well-defined nodal lines which divide the faces of the box. There are also 2 deflection shapes that are governed by the top part of the box. This behaviour is probably due to the different material of the top of the box which is made in plastic as reported in Table 3.

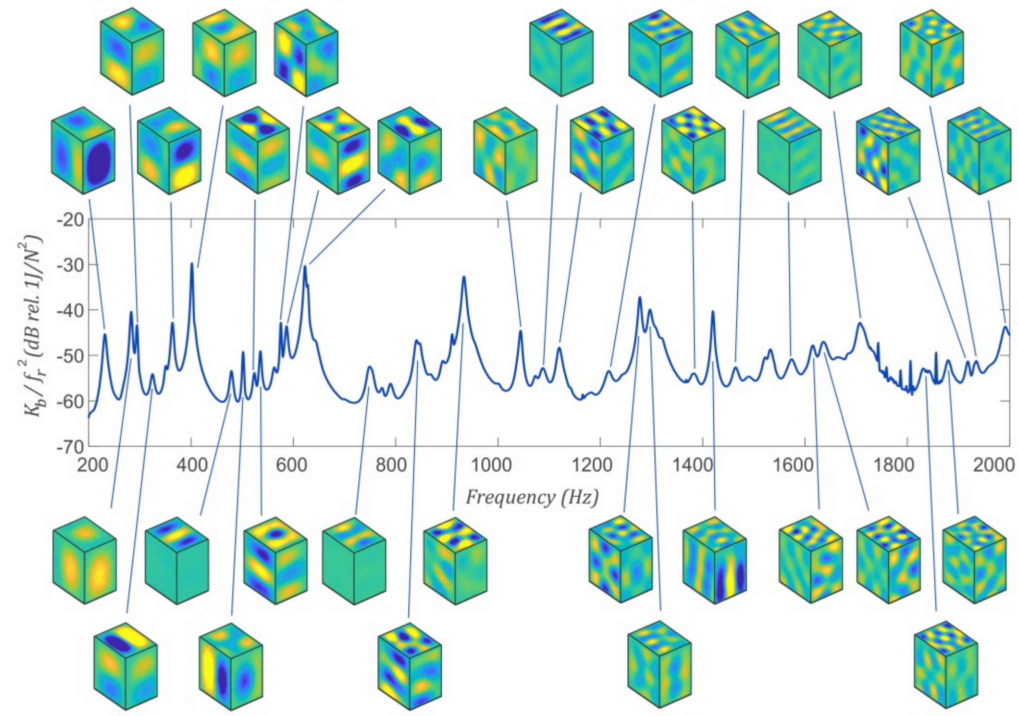


Figure 37: spectrum of the total flexural kinetic energy per unit force excitation derived from laser vibrometer measurements with the deflection shapes at the resonance frequencies.

Yet again, as previously done for the cylindrical structure, the results obtained from the laser vibrometer measurements are compared with those derived from high-speed camera data. Thus, in Figure 38, the orange curve represents the time-averaged total flexural kinetic energy per unit force excitation computed from the camera images, while the blue curve reports the corresponding results from the laser measurements. The agreement between the two spectra is overall quite good: most resonance frequencies are consistently identified by both methods, and the amplitudes of the resonance peaks generally show a comparable trend. As in the case of the cylindrical shell discussed above, minor discrepancies, including slight frequency shifts and differences in peak magnitudes, are present and can be attributed to the non-simultaneity of the measurements. In particular, the laser acquisition, which required several hours, was more susceptible to environmental fluctuations such as ambient temperature variations and internal heating effects caused by the shaker placed inside the structure. It is worth noting that, in analogy to what was observed in the case of the cylinder, the spectrum obtained from the camera data tends to exhibit slightly higher noise levels at both the lower and upper ends of the frequency range.

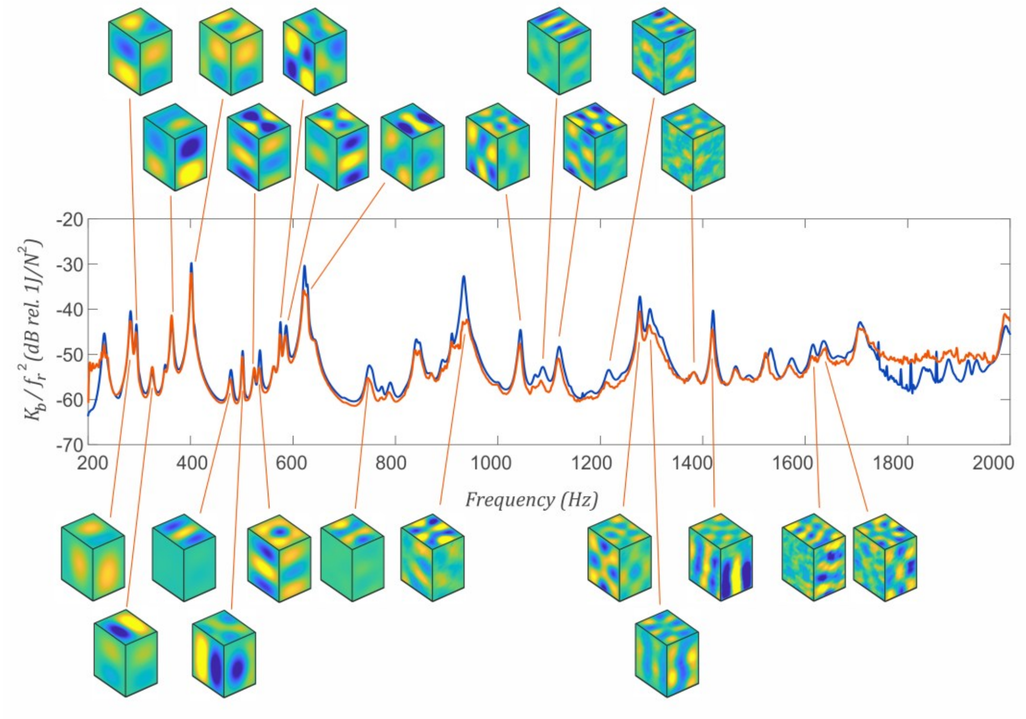


Figure 38: spectrum of the total flexural kinetic energy per unit force excitation derived from camera measurements (orange line) with the deflection shapes at the resonance frequencies (compared with laser vibrometer measurements, blue line).

Additionally, it should be pointed out that, in the case of the box, the analysed frequency band is broader than the one used for the cylinder. As a result, less excitation energy is available per frequency during the measurement process, leading to smaller structural deformations and consequently making the displacement identification via *DIC* more challenging. Nevertheless, the consistency between the two datasets remains strong in the majority of the frequency range considered, as further evidenced by the deflection shapes reported in Figures 37 and 38. These shapes, especially at higher resonance frequencies, appear another time somewhat more irregular in the camera-based results compared to those reconstructed from the laser data. This difference may be due again to localised variations in sensitivity or reconstruction accuracy across the different views. Still, similarly to the cylinder case, such discrepancies have only a marginal influence on the overall vibration behaviour and do not significantly affect the interpretation of the dynamic response or the estimation of acoustic radiation.

SUMMARY OF THE CHAPTER

This chapter addresses the measurement and reconstruction of structural vibrations using a single high-speed, frame-based optical camera. The chapter first introduces the test structures, namely a thin-walled baffled cylinder and a small box, together with the experimental setups involving a high-speed camera and a scanning laser vibrometer used as reference instrumentation. After presenting the acquisition procedures, the chapter develops a frequency-domain triangulation framework that enables the reconstruction of three-dimensional vibration fields from multiple camera viewpoints using short video recordings.

The theoretical background of structural dynamics relevant to the investigated geometries is then recalled, including the natural modes of cylindrical shells and thin plates, providing the basis for interpreting the measured vibration responses. The proposed methodology is subsequently applied to both test structures, allowing the reconstruction of full-field deflection shapes, and kinetic energy spectra. These results are systematically compared with reference measurements obtained from laser vibrometer, demonstrating the capability of the optical approach to capture both modal characteristics and broadband vibration behaviour.

Overall, the chapter establishes a robust experimental and methodological foundation for the optical measurement of structural vibrations using a single high-speed camera, which is later extended in subsequent chapters to the derivation of acoustic radiation and to the investigation of event-based vision systems as an alternative optical approach for structural vibration measurements.

6
BAFFLED CYLINDER SOUND RADIATION

This chapter presents the experimental and analytical investigation of the sound radiation from a baffled cylinder. The description begins with the test rig configurations, including the use of dedicated microphone arrays for both acoustic field reconstruction and sound power level spectra derivation. The derivation of the acoustic field is addressed through the discretised Kirchhoff–Helmholtz formulation [50,86], and the corresponding experimental results are reported. The sound power spectra are subsequently obtained using a discretised implementation in which the measurement surface is collapsed onto the vibrating surface of the cylinder, representing a non-standard yet effective approach. The chapter concludes with the presentation of results in both narrow-band and one-third-octave-band formats, providing a comprehensive assessment of the set of results aimed at validating the camera-based methodology for deriving the acoustic radiation of structures which have an analytical solution for the acoustic field.

6.1 TEST RIGS: MICROPHONE ARRAYS

The acoustic field and the sound power spectra reconstructed from camera measurements were validated against reference measurements obtained using array of microphones [50]. The measurements were carried out in a $500 - 1500\ Hz$ frequency range such that the acoustic wavelength spanned the range between $0.69\ m$ and $0.23\ m$.

MICROPHONES FOR ACOUSTIC FIELD DERIVATION

Figure 39 illustrates the baffled cylindrical model structure described in Section 5.1 with two additional planes used for acoustic field measurements and derivation. The sound field on the vertical plane was measured at the two grids of $12 \times 14 = 168$ points depicted in Figure 39a with the vertical array of microphones depicted in Figure 39b, which were suitably translated in radial direction to obtain the whole field. Also, the sound field on the horizontal measurement plane was measured at the grid of $24 \times 14 = 336$ points depicted in Figure 39c with the radial array of microphones shown in Figure 39d. Here, as done for the vibration measurements with the laser vibrometer, the microphones array was aligned along a radial segment and kept in the same position whereas the cylinder was rotated by $360\ deg$ to obtain the whole field. More specifically, the *FRFs* of the sound pressure per unit force excitation at each point in the outer acoustic domain were measured using the arrays of microphones described above. At the end, the maps of the acoustic field were derived and these maps depicted in great detail the lobes due to the near field sound radiation generated by the flexural vibration of the cylinder.

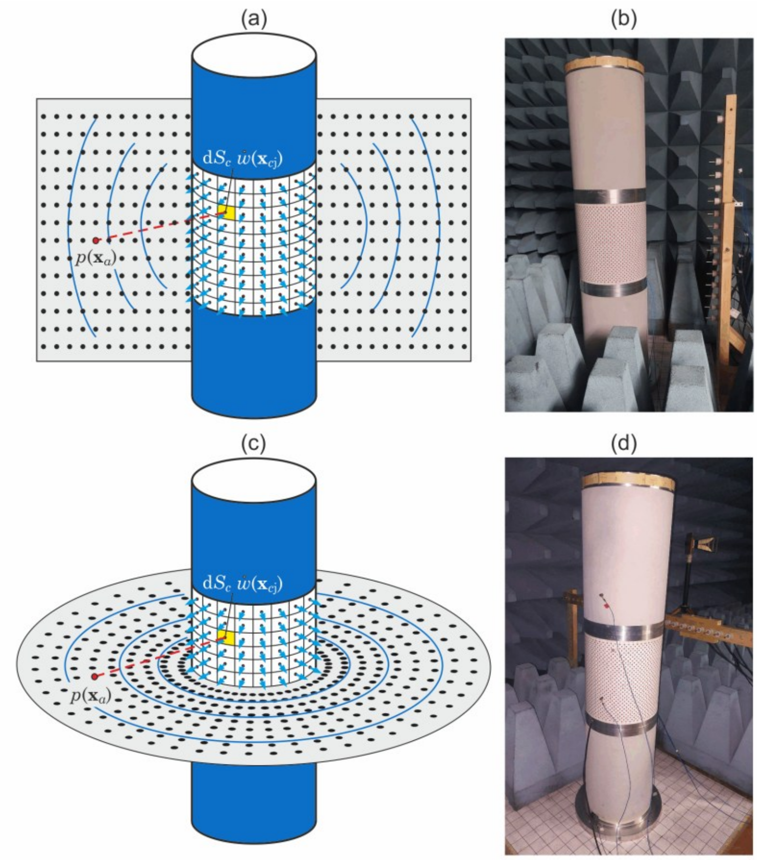


Figure 39: derivation of the sound fields in (a-b) vertical and (c-d) horizontal planes from the radiations of a mesh of quadrangular pistons in the rigid cylindrical baffle.

MICROPHONES FOR POWER LEVEL SPECTRA DERIVATION

The total sound radiation was measured with the vertical array of microphones shown in Figure 40. The measurement was carried out according to the *ISO 3740:2019* and *ISO 3744:2010* guidelines [52,54] for the estimation of sound radiation in an anechoic room.

In general, this measurement is carried out in a semi-anechoic room by placing the tested object on the floor and taking sound pressure measurements on a grid of points lying on a large hemispherical surface where the sound field produced by the radiator can be locally approximated as a plane wave such that the radiated sound power can be derived from measurements of the sound pressure only [10].

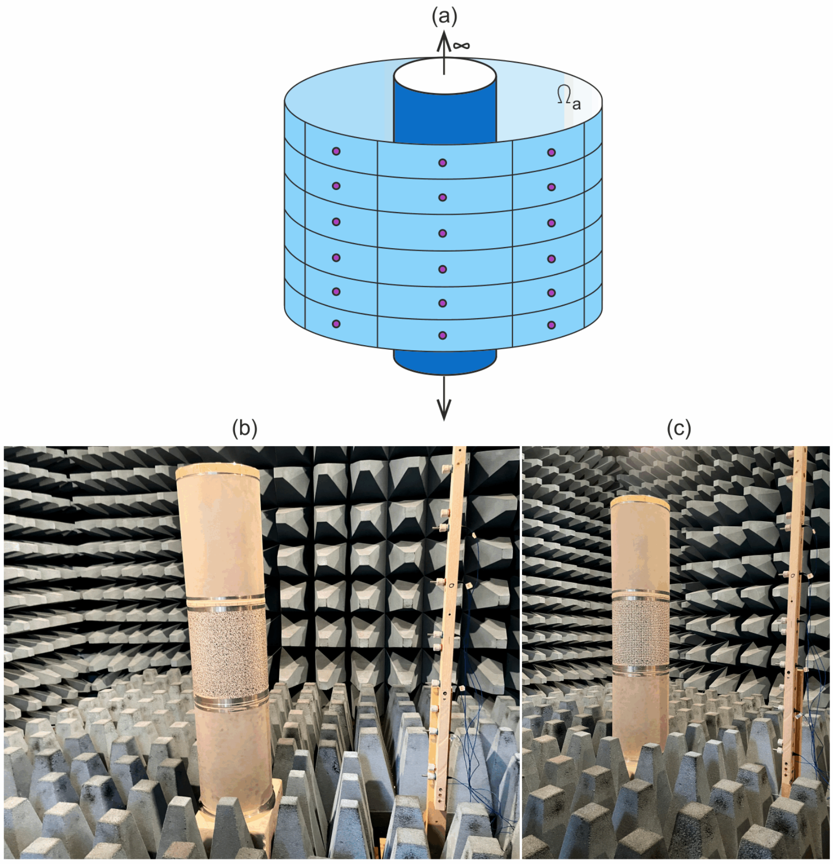


Figure 40: a) sketch of classical far-field sound radiation measurement surface with 8×6 mesh of quadrilateral elements and (b-c) measurements setup in the anechoic room

Given the peculiar shape of the baffled cylinder radiator considered here, as can be seen in Figure 40b, the measurement was carried out on an anechoic room and the sound pressures were taken on a large cylindrical surface, which was extended over the two baffles. In this way, both the direct sound radiation by the cylinder and the scattering effect produced by the two baffles were measured appropriately. For the given dimensions of the baffled cylinder and of the whole anechoic room, measurement cylindrical surface was taken to have a diameter $D_s = 1500\ mm$ and a height $H_s = 600\ mm$ such that it extends over the baffles by $\Delta H = 150\ mm$. The sound pressures were measured on an 8×6 grid of points (8 in circumferential and 6 in axial direction). The measurement points were located at the centre of an 8×6 regular mesh of quadrilateral elements. As can be noticed in Figure 40b, the measurements were taken with the vertical array of 6 microphones. As done with the laser vibrometer, here too, the acquisitions from the array of 6 microphones were repeated 8 times with $45\ deg$ rotations of the baffled cylinder.

6.2 KIRCHHOFF-HELMHOLTZ FORMULATION FOR ACOUSTIC FIELD DERIVATION

As described in Chapter 4, the sound radiated by a vibrating structure can be derived from the Kirchhoff–Helmholtz formulation, as expressed in Equation (4.29). As introduced in Section 5.1, the baffled cylinder examined in this thesis was chosen because the computation of the far-field sound pressure can be calculated in analytical form. More specifically, Equation (4.29) can be solved through a *single-step* procedure rather than a *two-step* approach. In the case of a baffled cylindrical shell, the surface integral can be evaluated analytically provided that a Green's function is constructed to satisfy the Neumann boundary condition, that is $\partial G(|\boldsymbol{x_{ai}} - \boldsymbol{x_c}|, \omega)/\partial r = 0$, on the surface of both the cylinder and the baffles (here $\boldsymbol{x_{ai}} = (r_i', \theta_i', z_i')$ denotes a point in the free field and $\boldsymbol{x_c} = (R, \theta, z)$ represents a point on the surface of the cylinder). As discussed in [56,84,100], this condition implies that the acoustic field generated by an elemental source is scattered by the structural boundaries as if they were rigid. Under these assumptions, the unknown boundary pressure term $p(\boldsymbol{x_c}, \omega)$ vanishes, allowing the Kirchhoff–Helmholtz integral to be simplified into the following expression:

$$p(\boldsymbol{x_{ai}}, \omega) = -i\omega\rho_0 \int_{S_c} \big(G(|\boldsymbol{x_{ai}} - \boldsymbol{x_c}|, \omega)\dot{u}_r(\boldsymbol{x_c}, \omega)\big)\, dS_c. \tag{6.1}$$

In general, the analytical evaluation of this integral is feasible when the boundary surface can be described using a single coordinate of an appropriate coordinate system, such as cylindrical coordinates in the present case, and when the acoustic wave equation is separable within that coordinate framework [56]. This condition is met by the baffled cylindrical structure examined in this study, for which the following analytical solution can be formulated in cylindrical coordinates [101–103]:

$$p(\boldsymbol{x_{ai}}, \omega) = -\frac{i\omega\rho_0}{4\pi^2} \int_0^{2\pi} \int_0^{h_c} \dot{u}_r(R, \theta, z, \omega) \sum_{n=-\infty}^{+\infty} \cos\big(n(\theta_i' - \theta)\big) I_n(z_i' - z)\, d\theta dz, \tag{6.2}$$

where:

$$I_n(z_i' - z) = \int_{-\infty}^{+\infty} \frac{\cos\big(k_z(z_i' - z)\big)}{k_r R} \frac{H_n^{(2)}(k_r r_i')}{H_n^{(2)'}(k_r R)} dk_z. \tag{6.3}$$

In these equations k_z is the projection into the longitudinal direction of the cylinder of the acoustic wavenumber $k = \omega/c$, where c is the speed of sound in air and $k_r = \sqrt{k^2 - k_z^2}$. Finally, $H_n^{(2)}$ and $H_n^{(2)'}$ are the e *2nd-kind* Hankel function and its derivative respectively [104]. Equation (6.3) holds for time dependence taken in the complex form $e^{i\omega t}$, if $e^{-i\omega t}$ was assumed to describe the co-sinusoidal time-harmonic vibration, then Equation (6.2) would have had in front a term $+i$ and would have encompassed the *1st-kind* Hankel function and therein derivative [104].

Since the radial velocity is reconstructed with the camera measurements at the grid of points shown in Figure 23a, the surface integral in Equation (6.2) has been calculated numerically with a middle Riemann sum considering the mesh of quadrangular radiators defined on the surface of the cylinder such that:

$$p(\boldsymbol{x_{ai}}, \omega) \cong -\frac{i\omega\rho_0 S_c}{4N\pi^2} \sum_{j=1}^{N} \dot{u}_r(\boldsymbol{x_{cj}}, \omega) \sum_{n=-\infty}^{+\infty} \cos\big(n(\theta_i' - \theta)\big) I_n\big(z_i' - z_j\big). \tag{6.4}$$

In this equation, S_c/N indicates the area of a single surface element, whose centres are identified by the coordinates (R, θ_j, z_j). After a few mathematical steps, this equation can be further simplified into the following expression:

$$\begin{aligned} p(\boldsymbol{x_{ai}}, \omega) &\cong -\frac{i\omega\rho_0 S_c}{4N\pi^2} \sum_{j=1}^{N} \dot{u}_r(\boldsymbol{x_{cj}}, \omega) \sum_{n=0}^{N'} \varepsilon_n \cos\big(n(\theta_i' - \theta)\big) \bar{I}_n\big(z_i' - z_j\big) \\ &= Z_{ij}(\omega) \dot{u}_r(\boldsymbol{x_{cj}}, \omega), \end{aligned} \tag{6.5}$$

where $\varepsilon_n = 1$ for $n = 0$, $\varepsilon_n = 2$ for $n > 0$, and the infinite summation is approximated into the sum of the first $N + 1$ terms $\bar{I}_n$. Here,

$$\bar{I}_n\big(z_i' - z_j\big) = \int_{0}^{+\infty} \frac{\cos\big(k_z(z_i' - z)\big)}{k_r R} \frac{H_n^{(2)}(k_r r_i')}{H_n^{(2)'}(k_r R)} dk_z. \tag{6.6}$$

It is important to notice that when $k_z > k$, $k_r = \sqrt{k^2 - k_z^2}$ become imaginary, and, in this case, it was set:

$$\frac{H_n^{(2)}(k_r r_i')}{k_r R H_n^{(2)'}(k_r R)} = \frac{K_n(|k_r| r_i')}{|k_r| R K_n'(|k_r| R)} \tag{6.7}$$

where K_n represents the modified Bessel function of the second kind.

The integral in Equation (6.6) was approximated into a middle Reimann sum too, such that:

$$\overline{I_n}(z_i' - z_j) = \sum_{m=0}^{M} \frac{\cos\left(m\Delta k_z(z_i' - z_j)\right)}{\sqrt{k^2 - (m\Delta k_z)^2} R} \frac{H_n^{(2)}\left(\sqrt{k^2 - (m\Delta k_z)^2} r_i'\right)}{H_n^{(2)'}\left(\sqrt{k^2 - (m\Delta k_z)^2} R\right)} \Delta k_z. \tag{6.8}$$

Here Δk_z and M were selected to achieve a reasonable approximation of the integral.

Now recalling that $\dot{u}_r(\boldsymbol{x_{cj}}, \omega) = i\omega u_r(\boldsymbol{x_{cj}}, \omega)$ and using Equation (5.30) into Equation (6.5) gives:

$$p(\boldsymbol{x_{ai}}, \omega) = i\omega Z_{ij}(\omega) \alpha_{rj}(\omega) f_r(\boldsymbol{x_f}, \omega). \tag{6.9}$$

In conclusion, for time-harmonic vibrations, the sound radiation field of the cylinder at given frequencies ω can also be reconstructed and depicted in terms of "operational acoustic shapes", which can be derived from Equation (4.18) rewritten in the following trigonometric form:

$$p(\boldsymbol{x_{ai}}, t) = P(\boldsymbol{x_{ai}}, \omega) \cos\left(\omega t + \varphi_P(\boldsymbol{x_{ai}}, \omega)\right), \tag{6.10}$$

where $P(\boldsymbol{x_{ai}}, \omega) = |p(\boldsymbol{x_{ai}}, \omega)| = \left|i\omega Z_{ij}(\omega) \alpha_{rj}(\omega) f_r(\boldsymbol{x_f}, \omega)\right|$ and $\varphi_P(\boldsymbol{x_{ai}}, \omega) = \angle\{p(\boldsymbol{x_{ai}}, \omega)\} = \angle\left\{i\omega Z_{ij}(\omega) \alpha_{rj}(\omega) f_r(\boldsymbol{x_f}, \omega)\right\}$ are respectively the amplitude and phase of the sound pressure at each point where the sound field is depicted. In this thesis, the operational acoustic shapes were derived from Equation (6.10) with respect to the same instant of time of maximum amplitude of the flexural vibration fields. Equations (6.9) and (6.10) indicate that, for time-harmonic vibrations, the amplitude of the acoustic pressure at a given instant of time is modulated by the structural response of the thin-walled cylinder, i.e. via the dynamic flexibility $\alpha_{rj}(\omega)$, and by the acoustic response generated by the radiation of the baffled cylinder, i.e. via the acoustic impedance $Z_{ij}(\omega)$.

Moving now to the results obtained with the procedure described above, the sound radiation is instead assessed at the monitor position $r_m' = 219\ mm$, $\theta_m' = 0\ deg$ and $z_m' = 0\ mm$ with respect to the time-averaged sound pressure per unit force excitation either measured with a microphone and shown in Figure 41 using a violet line, or derived from Equation (6.9) using the vibration *FRFs* acquired with the laser vibrometer (blue line in Figure 41a) or reconstructed from the camera measurements (orange line in

Figure 41b). Comparing first the spectrum of the sound pressure measured with the microphone with the spectrum in Figure 34 for the spatially averaged vibration measured with the laser vibrometer it is noted that the two spectra show similar features, although the resonance peaks in the acoustic spectrum have quite an uneven range of amplitudes and there are some antiresonance troughs in between resonances. This is the result of the fact that, at each frequency, the amplitude of the acoustic field is modulated by the radiation efficiency of the cylinder flexural modes that mainly contribute to the response at that frequency [58]. Therefore, the sound pressure at certain resonance frequencies, where the response is controlled by the resonant mode [58,97], may be significantly attenuated when the resonant mode has a low radiation efficiency. Moreover, the acoustic spectrum provides the sound pressure at a point where at certain frequencies the sound field may be characterised by destructive interference effects that significantly reduce the sound pressure. The plots in Figure 41 show that, in general, the spectra of the sound pressure derived from the laser vibrometer (blue line in Figure 41a) or the camera (orange line in Figure 41b) measurements reproduce rather well the principal features of the spectrum measured directly with the microphone.

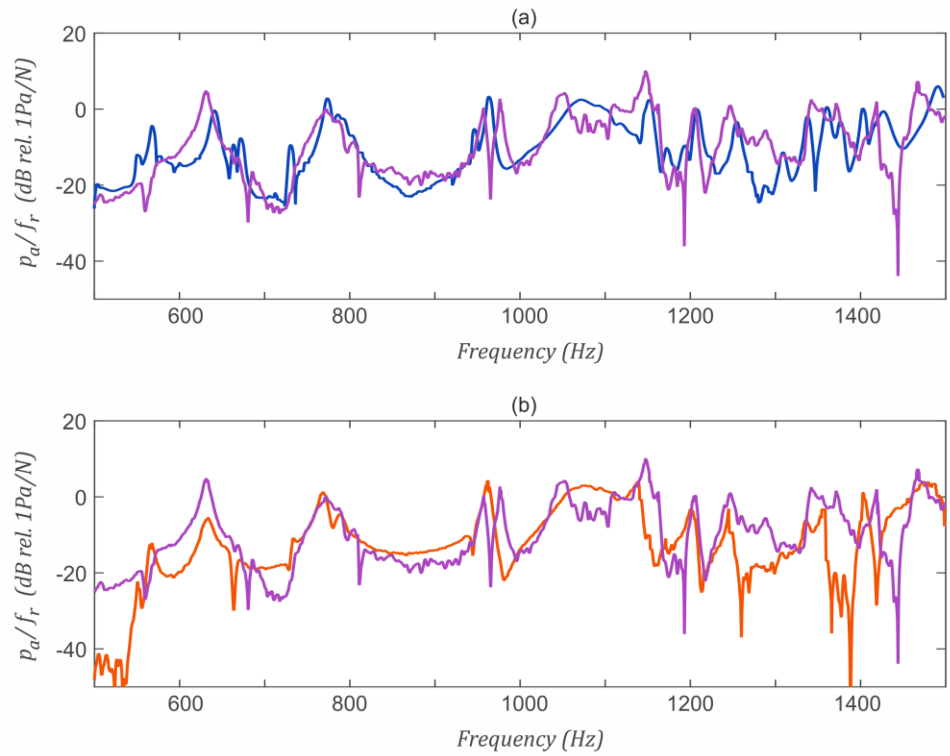


Figure 41: spectra of the sound pressure at a monitor point per unit force excitation derived from microphone measurement (violet lines), laser vibrometer measurements (blue line) and camera measurements (orange line).

There are some marked differences over small frequency bandwidths, whose nature is difficult to identify. In fact, the spectra based on direct measurements of the sound pressure may have been affected by small noise (e.g. the acquisition system that was kept in the measurement room) and flanking sources (e.g. structure-borne noise due to vibrations transmitted from the cylinder to the floor of the anechoic room). Instead, the spectrum derived from the laser vibrometer measurements may have been affected by the variability of the flexural response of the cylinder during the long-lasting measurements (several hours). Finally, the spectrum derived from the camera measurements may have been limited by the spatial resolution of the camera employed in this work. Rather importantly, the three set of measurements couldn't be taken simultaneously and thus the response of the structure may have changed from one measurement to another. Nevertheless, the spectra shown in Figure 41 indicate that the proposed approach based on the reconstruction of the sound radiation using a discretized version of the Kirchhoff-Helmholtz integral and the vibration field reconstructed from multi-view measurements taken with an optical camera can be successfully employed to derive the sound radiation from a closed shell over a wide audio frequency range.

As anticipated in Section 6.1, the maps of the acoustic field were derived using both the *FRFs* obtained from microphones measurements and the *FRFs* derived from laser and camera flexural vibration data. Starting from laser and camera data, the *FRFs* of the sound pressure per unit force excitation at each point in the outer acoustic domain were derived with a midpoint-Riemann sum of the Kirchhoff – Helmholtz integral based on the mesh of quadrangular elements depicted in Figure 23a. Considering the physics of the sound radiation phenomenon, this corresponds to accounting for the superposition of the sound radiated by the radial vibration of the mesh of quadrangular elements, assuming each element vibrates in a rigid cylindrical baffle of infinite length [56,58]. The sound radiation fields per unit force excitation were then reconstructed from the vibration and sound radiation *FRFs* and represented considering a fixed instant of time in which the amplitude of the flexural vibrational field of the cylinder is maximum. The time-harmonic vibration and acoustic fields were derived at four resonance frequencies of the cylinder flexural vibration, that is at $f_1 = 618\ Hz$, $f_2 = 659\ Hz$, $f_3 = 964\ Hz$, $f_4 = 1231\ Hz$. Hence, considering a time instant where the amplitude of the vibration is maximum, the flexural field closely reproduces the shape of the resonant mode, which is normally called 'Operating Deflection Shape' (*ODS*) [8]. Likewise, the acoustic field radiated in the proximity of the cylinder, i.e. the acoustic nearfield, is characterised by distinct lobes that replicate the circular shape of the flexural mode that controls the resonant response, that is the operating deflection shape, at the given resonance frequency. For the resonance frequency $f_4 = 1231\ Hz$, the map was reconstructed at $3/4$ of the cylinder height since the vibration field was characterised by a nodal line at the midspan of the cylinder. In fact, given the symmetry of the vibration field, the sound radiation at the *mid-plane* would show a very small uniform distribution of the sound pressure due to mutual cancellation effects from the anti-symmetric radiations produced by the vibrations of the top and bottom half of the cylinder.

The spectrum of the cylinder flexural vibration is characterised by sharp peaks (as described in Chapter 5), which suggest that at most resonance frequencies, the response is controlled by the resonant mode only, it is expected that the radiated sound field near the cylinder surface should be characterised by a circle of acoustic lobes that comply with the circumferential lobes of the resonating flexural mode. The reconstructed and measured acoustic fields are shown in Figure 42, Figure 43, Figure 44, Figure 45. More specifically, the maps (a) show the vibration and acoustic fields reconstructed from camera measurements. The maps (b) show the vibration and acoustic fields obtained directly from, and reconstructed from, the laser vibrometer measurements. Finally, the maps (c) show the fields obtained from direct measurements with the laser vibrometer and the microphones array respectively.

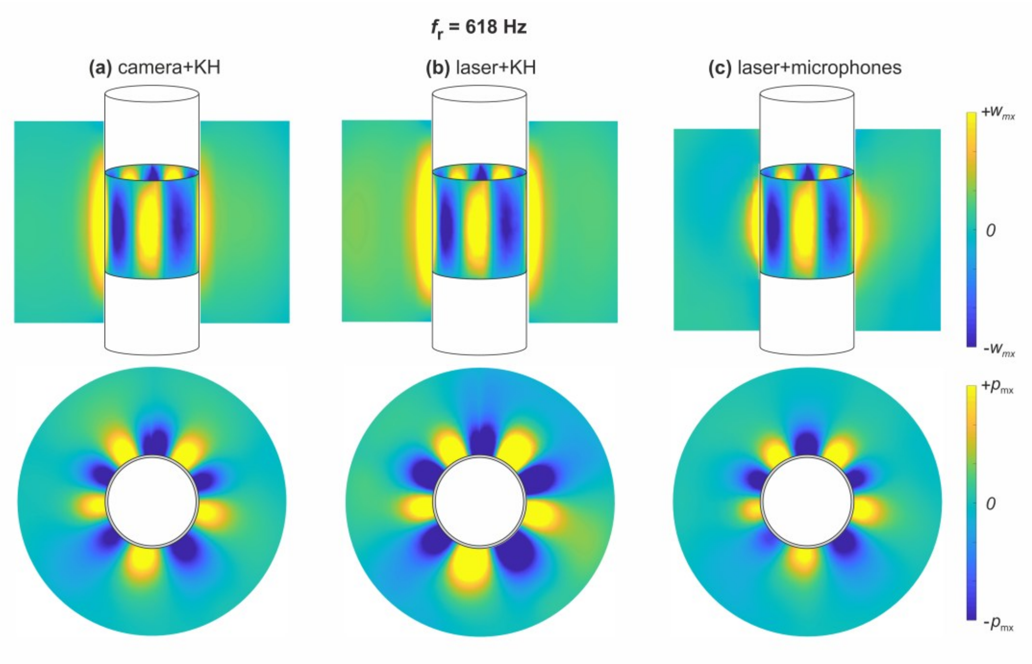


Figure 42: vibration and sound radiation fields reconstructed from (a) camera measurements and KH simulation, (b) laser-vib. measurements and KH simulation, (c) laser-vib. measurements and microphones-array measurements at the resonance frequency $fr = 618\ Hz$.

Figure 42 shows that the time-harmonic vibration field at $f_1 = 618\ Hz$ is characterised by the natural flexural mode with axial mode order $n_z = 1$ and circumferential mode order $n_\theta = 5$. The field reconstructed from the camera measurement depicted in (a) closely overlaps with that measured with the laser vibrometer shown in (b) and (c). The magnitude of the field derived from the laser vibrometer measurements is slightly bigger than that obtained from the camera measurements. This mismatch is however marginal and possibly due to the fact that the two measurements could not be taken simultaneously, primarily because the camera acquisitions took much less time than the measurements with the laser vibrometer. The acoustic map in (a) obtained from the camera measurements shows that, in the vicinity of the cylinder, the radiated acoustic field in the horizontal midplane is characterised by 10 lobes, which closely replicates the wavy shape of the flexural mode in circumferential direction. The same effect is seen in the vertical plane where the acoustic field is characterised by a single lobe that matches the vibration field in vertical direction. The acoustic field (a) reconstructed from the camera measurements agrees quite well both with that reconstructed from the laser vibrometer measurements shown in (b) and with that measured directly with the microphone array depicted in (c). The nearfield lobes in (b) extend over a slightly bigger area, but this may be the result of the small mismatch in the amplitude of the camera and laser vibrometer vibration fields discussed above. Instead, the lobes in (c) extend over a slightly smaller surface, and this could be the result of the approximations introduced by the measurement setup. In general, the small differences between the reconstructed and measured vibration and sound fields could simply be due to the fact that the measurements cannot be taken simultaneously and thus the response of the cylinder as well as the excitation generated by the shaker may have changed a little from measurement to measurement.

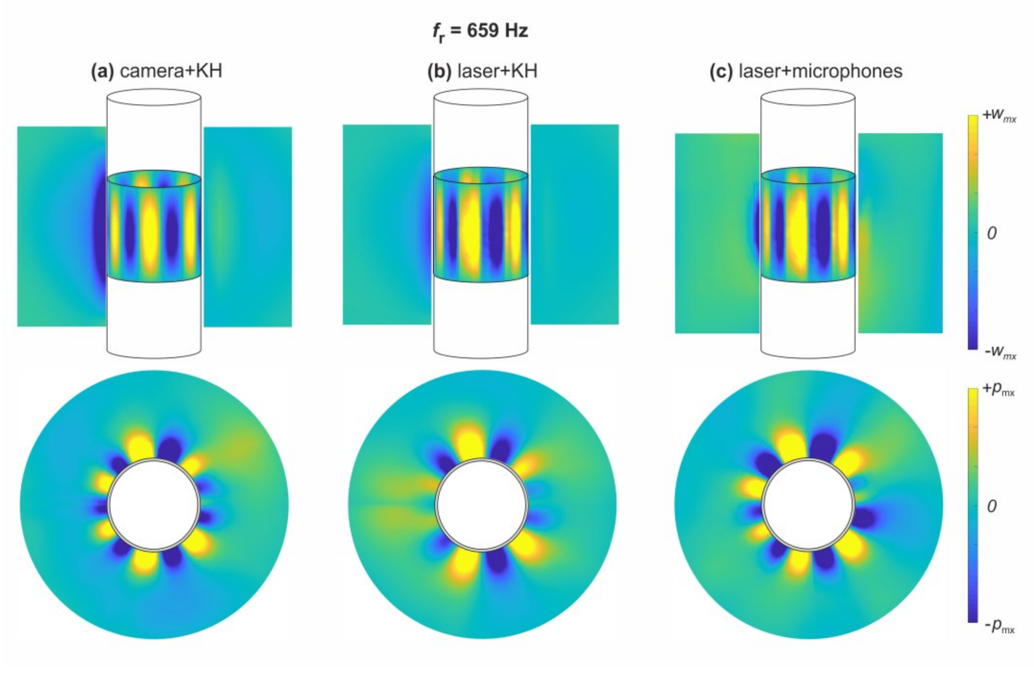


Figure 43: vibration and sound radiation fields reconstructed from (a) camera measurements and KH simulation, (b) laser-vib. measurements and KH simulation, (c) laser-vib. measurements and microphones-array measurements at the resonance frequency $fr = 659\ Hz$.

Moving to the time-harmonic excitation at $f_2 = 659\ Hz$, and considering Figure 43, it is notable that the vibration field is controlled by the natural flexural mode with axial mode order $n_z = 1$ and circumferential mode order $n_\theta = 7$. Here too, contrasting the three vibration maps, the field reconstructed from the camera measurement in (a) overlaps quite well with that measured with the laser vibrometer in (b) and (c). The acoustic nearfield reconstructed from the camera measurements shown in (a) is characterised by 14 lobes with alternating *positive–negative* sound pressure in the horizontal mid plane and a single lobe in the vertical plane, which copy the wavy shape of the flexural mode in circumferential and axial directions. The acoustic field (a) reconstructed from the camera measurements agrees rather well with that measured directly with the microphones array depicted in (c) apart from small mismatches, which, in this case too, are likely to be due to the fact that the two sets of measurements could not be taken simultaneously. Also, the acoustic field (a) closely replicates that reconstructed from laser vibrometer measurements shown in (b) apart from the lobes on the left-hand side, which, however, were not properly reconstructed from the laser vibration measurements that presented small errors in that area.

Considering now Figure 44, the vibration measurements show that the time-harmonic vibration field at $f_3 = 964\ Hz$ is controlled by the natural flexural mode with axial mode order $n_z = 1$ and circumferential mode order $n_\theta = 9$. Again, the field in (a) obtained from the camera measurements agrees well with that in (b) and (c) measured directly with the laser vibrometer. Here, the shape of the radiated acoustic field is somewhat more complex than that found at the first two resonance frequencies. Once more, the field in the horizontal plane presents a ring of lobes with alternating *positive–negative* sound pressure that somehow replicates the wavy vibration field in the circumferential direction of the cylinder. However, in this case, there is also a second ring formed by a total of 6 bigger lobes, yet with alternating *positive–negative* sound pressures. These features are found too in the acoustic fields reconstructed from the laser vibrometer measurements and in those measured directly with the array of microphones. Here, the shape and the intensity of the small lobes adjacent to the cylinder are not the same for all lobes. This is probably the result of two concurrent effects. On the one hand the vibration field with high circumferential mode order generates these smaller lobes, whose shape and intensity can be affected primarily by small errors in the measurements. Also, the longitudinal welding of the cylinder wall may have had a bigger impact on the sound radiation. These features reflect quite significantly on the lobe in the vertical plane too. In fact, the small deviations in circumferential directions of the nearfield lobes, generate quite significant differences in the sound fields depicted in the vertical planes. This mismatch could be due to the fact that the acoustic measurements on the vertical plane were taken with a small angular offset too.

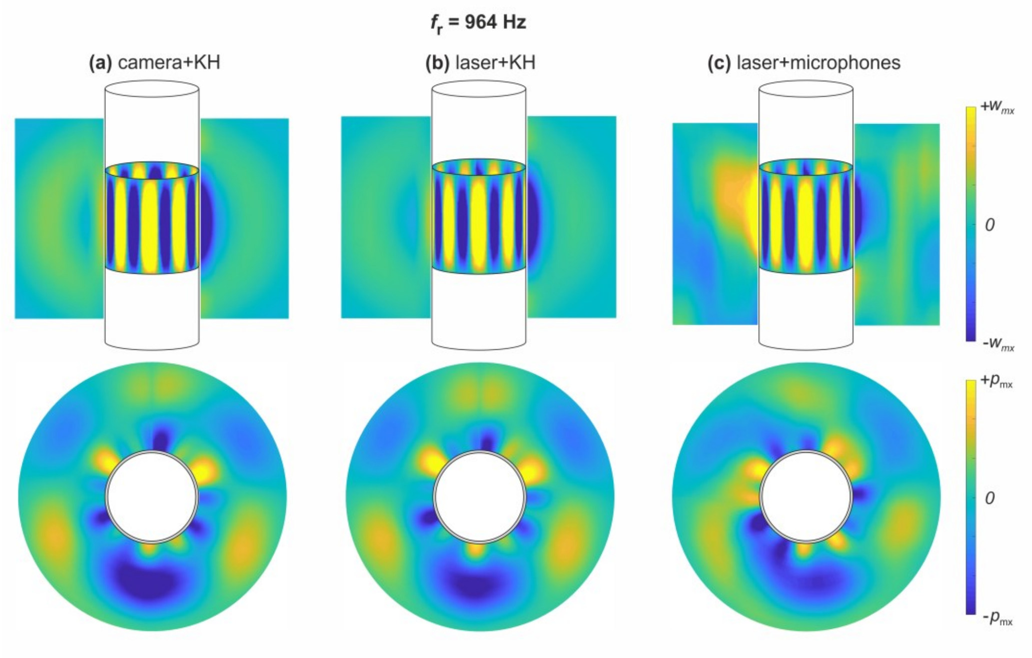


Figure 44: vibration and sound radiation fields reconstructed from (a) camera measurements and KH simulation, (b) laser-vib. measurements and KH simulation, (c) laser-vib. measurements and microphones-array measurements at the resonance frequency $fr = 964\ Hz$.

Overall, the results presented in Figure 44 confirm a good agreement of the vibration and sound radiation fields reconstructed from the camera measurements with both those obtained from the laser vibrometer measurements and those measured directly with the laser vibrometer and microphone array.

To conclude this section, the maps depicted in Figure 45 for the flexural vibration and sound radiation at $f_3 = 1231\ Hz$ are analysed here. As can be seen in the top pictures, in this case the flexural vibration field is characterised by a flexural mode with axial mode order $n_z = 2$ and circumferential mode order $n_\theta = 6$. Hence, the vibration field is cut in two sections along the length, with a circular nodal line at the middle span, which divides the alternating vibration lobes in two rings having an angular offset such that the $n_z = 2$ positive and negative vibration lobes develop along the length. Here the vibration field (a) reconstructed from the camera measurements is slightly irregular, although, overall, it reproduces the shape obtained from the measurement taken with the laser vibrometer (b) and (c). This is likely to be caused by the fact that, as can be noticed in the spectrum plotted in Figure 35, the measurements are close to the upper limit of the frequency range that can be handled by the camera.

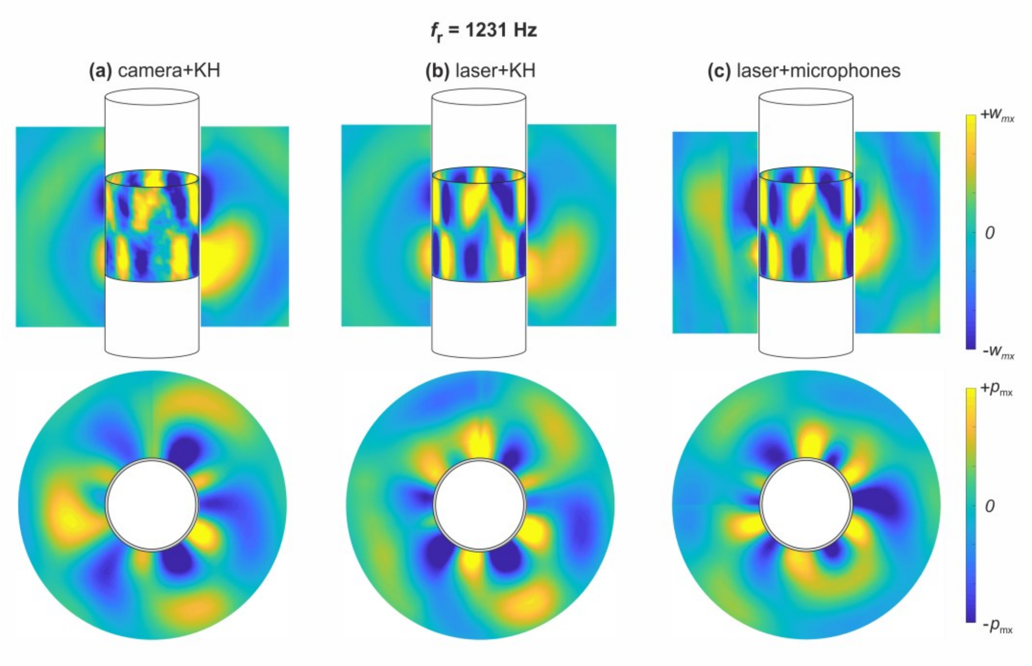


Figure 45: vibration and sound radiation fields reconstructed from (a) camera measurements and KH simulation, (b) laser-vib. measurements and KH simulation, (c) laser-vib. measurements and microphones-array measurements at the resonance frequency $fr = 1231\ Hz$.

Nevertheless, as can be seen in the map (a), the sound field reconstructed from the camera acquisitions, is quite smooth and, in fact, it replicates reasonably well both that reconstructed from the laser vibrometer measurement in (b) and that measured directly with the arrays of microphones in (c). Here again the acoustic field is characterised by two rings of lobes having alternated *positive–negative* sound pressures. The inner ring of lobes conforms to the wavy flexural vibration with circumferential mode order $n_\theta = 6$. The sound field in the vertical plane reconstructed from the camera acquisitions (a) shows the two *positive–negative* sound pressure lobes that conform to the vibration field in axial direction. The field replicates quite well that generated from the vibration measurements taken with the vibrometer as well as that measured directly with the microphones. Overall, despite the vibration field reconstructed from the camera measurements showing quite an irregular shape, the resulting acoustic field calculated with the approximated Kirchhoff-Helmholtz formulation is rather smooth and, also, agrees reasonably well with those either reconstructed from the laser vibrometer measurements or measured directly with the microphone array. This behaviour has been observed to stem from the presence of an irregular vibration field, which is influenced by a random distribution of small errors introduced during the digitalization process within the camera. However, since the acoustic pressure at each location in the free field arises from the combined radiation emitted by the entire surface of the cylinder, these errors tend to accumulate in a manner that leads to their mutual cancellation. As a result, a reasonably accurate representation of the radiated sound field can still be obtained.

To conclude, the accuracy of the reconstructed vibration and acoustic fields has been assessed quantitatively by evaluating the differences in decibels, interpreted as logarithmic ratios in linear scale, between the modulus-squared quantities of reconstructed and measured fields. Specifically, the comparison involves the summation of the squared magnitudes of the complex radial velocities and sound pressures, computed using the following expressions:

$$\begin{cases} \Delta \dot{u}_r(\omega_{res}) = 10\log_{10}\dfrac{\sum_j\left|\dot{u}_r^{cam}(\boldsymbol{x}_j,\omega_r)\right|^2}{\sum_j\left|\dot{u}_r^{vib}(\boldsymbol{x}_j,\omega_r)\right|^2} \\ \Delta p_{c-m}(\omega_{res}) = 10\log_{10}\dfrac{\sum_i|p^{cam}(\boldsymbol{x}_{ai},\omega_r)|^2}{\sum_i|p^{mic}(\boldsymbol{x}_{ai},\omega_r)|^2} \\ \Delta p_{v-m}(\omega_{res}) = 10\log_{10}\dfrac{\sum_i|p^{vib}(\boldsymbol{x}_{ai},\omega_r)|^2}{\sum_i|p^{mic}(\boldsymbol{x}_{ai},\omega_r)|^2} \\ \Delta p_{c-v}(\omega_{res}) = 10\log_{10}\dfrac{\sum_i|p^{cam}(\boldsymbol{x}_{ai},\omega_r)|^2}{\sum_i|p^{vib}(\boldsymbol{x}_{ai},\omega_r)|^2} \end{cases} \quad (6.11)$$

Here, $\dot{u}_r^{cam}(\boldsymbol{x}_j,\omega_r)$ and $\dot{u}_r^{vib}(\boldsymbol{x}_j,\omega_r)$ denote, respectively, the complex radial velocity reconstructed from the *camera-based* measurements and that obtained via laser vibrometer at the $j-th$ point on the vibration grid (as illustrated in Figure 23a). Similarly, $p^{cam}(\boldsymbol{x}_{ai},\omega_r)$, $p^{vib}(\boldsymbol{x}_{ai},\omega_r)$ and $p^{mic}(\boldsymbol{x}_{ai},\omega_r)$ correspond to the complex sound pressures at the $i-th$ point on the acoustic grids (shown in Figure 39a,c), as reconstructed from camera and laser vibrometer measurements, and as directly measured by the microphone array. The results reported in Table 4 reveal that, on average, the discrepancy between the vibration field reconstructed from the camera recordings and that acquired via laser vibrometer falls within the range of 0.7 to 3.2 dB. The differences are all positive, which suggests there could have been a systematic error, possibly due to the fact that the two measurements were taken not simultaneously and, actually, the laser vibration measurement lasted few hours compared to the few tens of minutes necessary for the camera acquisitions. Despite this limitation, the observed discrepancies confirm the viability of *camera-based* optical methods for capturing vibrational behaviour, even at relatively high audio frequencies. Moving to the acoustic field reconstruction, the data in Table 4 show that the sound field reconstructed from the camera measurements differs from that directly measured by the microphone array by an average of 0.4 to $2\,dB$. Furthermore, the discrepancy between the camera-reconstructed field and the one reconstructed from laser vibrometer measurements ranges from 0.6 to $4\ dB$. It is also worth noting that there are some differences between the two sets of errors, which are confirmed by the 0.2 to $4.2\ dB$ differences between the sound field reconstructed from the laser vibrometer measurements and measured directly with the microphones array.

Table 4: Differences in the spatially averaged cylinder radial velocities estimated from the camera measurements and measured with the laser vibrometer and differences of the spatially averaged radiated sound pressure reconstructed from camera measurements and laser vibrometer measurements and measured with microphones at the four reference resonance frequencies.

f_n (Hz)	Cylinder	Vertical Plane Δp (dB)			Horizontal Plane Δp (dB)		
	$\Delta\dot{u}_r$ (dB)	Δp_{c-m}	Δp_{c-v}	Δp_{v-m}	Δp_{c-m}	Δp_{c-v}	Δp_{v-m}
618	$+0.7$	-2	$+0.1$	-2	-2.5	-0.6	-1.9
659	$+2.1$	-1.9	$+1$	-2.9	-1.8	$+1.3$	-3.2
964	$+3.2$	$+0.4$	$+4.6$	-4.2	$+0.4$	$+4.4$	-4.1
1231	$+2.7$	$+1.2$	$+2$	-0.8	$+1$	$+1.2$	-0.2

There is no a clear correlation between the three sets of errors, which, yet again, may be the result of the fact that the three sets of measurements could not be implemented simultaneously and more importantly lasted for significantly different amounts of time. Overall, the 0.4 to $2\ dB$ errors found between the fields reconstructed with the cameras and those measured directly with the microphone array suggest that the proposed methodology offers a promising and practical approach to the reconstruction and analysis of radiated sound fields.

6.3 SOUND POWER LEVEL SPECTRA DERIVATION

In order to compute the sound power level, all experiments on the cylindrical structure were repeated to acquire updated surface vibration data using both a laser vibrometer and a high-speed camera, as well as new acoustic measurements at standardized *ISO* microphone positions. This repetition was necessary because the structure, having undergone a period of rest, disassembly, transport, and reassembly, may have experienced slight modifications, potentially affecting its dynamic behaviour. Although the newly acquired kinetic energy data is not reported again in detail it is briefly plotted in Figure 46 to illustrate that the differences with respect to the data presented in Chapter 5 are minimal and within an acceptable range of variation. Additionally, this new test campaign allowed for a slight extension of the measured frequency range, which was increased from $400\ Hz$ to $1600\ Hz$.

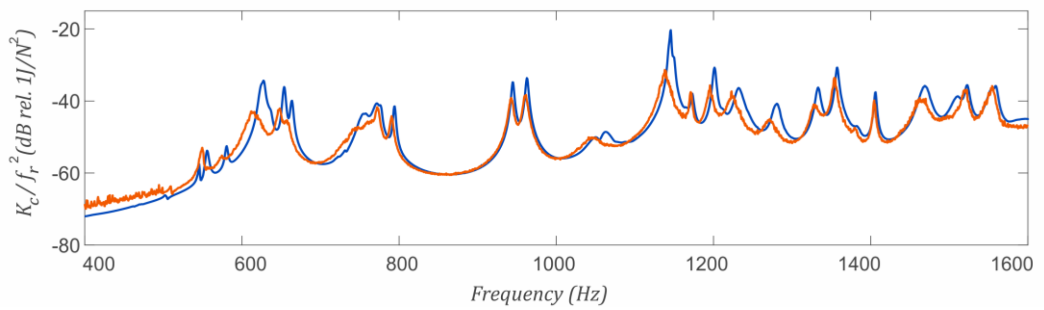


Figure 46: narrow band spectra of the total vibration kinetic energy per unit force excitation from measurements with laser vibrometer (blue line) and camera (orange line).

As described in Chapter 4, the sound power radiated by a vibrating structure can be derived from Equation (4.49), after having derived the pressures on the surface of the cylinder using the Kirchhoff-Helmholtz formulation discretized in Section (6.2). In this study a simpler approach is proposed, where, as depicted in Figure 47a,b, the closed surface Ω_a is actually collapsed onto the surface of the radiating structure, that is on the surface of the flexible cylinder S_c.

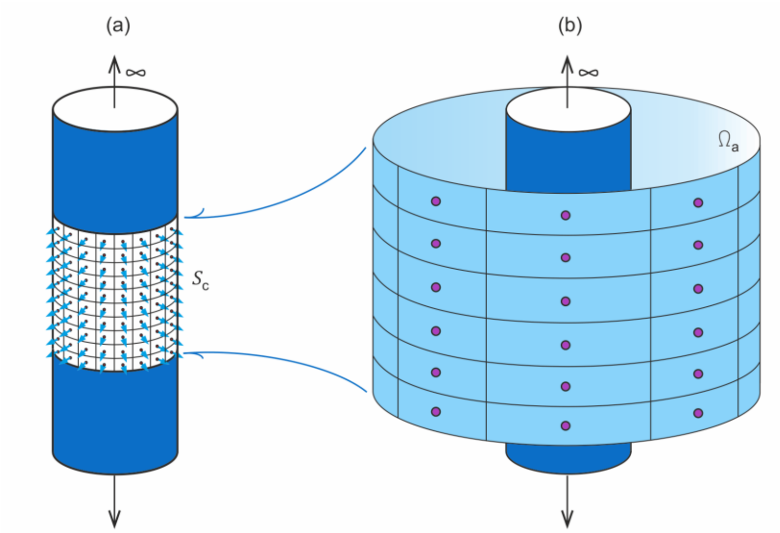


Figure 47: c) proposed near-field sound radiation measurement surface with 62×20 mesh of quadrilateral elements and d) classical far-field sound radiation measurement surface with 8×6 mesh of quadrilateral elements.

In this case, the fluid particle velocity over the surface of the radiating structure coincides with the radial, i.e. flexural, velocity of the cylinder, such that the instantaneous total sound power radiation results given by:

$$P(\omega) = \frac{1}{2} \int_{S_c} \mathcal{R}e\{p^*(\boldsymbol{x_c}, \omega)\dot{u}_r(\boldsymbol{x_c}, \omega)\} dS_c, \tag{6.12}$$

where $(\quad)^*$ is the complex-conjugate of the argument.

Considering the Kirchhoff-Helmholtz formulation and substituting the value of $p(\boldsymbol{x}, \omega)$ into (6.12), the time averaged total sound power radiation is given by the following double integral:

$$\bar{P}(\omega) = \frac{1}{2} \int_{S_c} \int_{S_c} \mathcal{R}e\{\dot{u}_r(\boldsymbol{x_c}, \omega)^* Z(\boldsymbol{x_c}|\boldsymbol{x_a}, \omega)^* \dot{u}_r(\boldsymbol{x_a}, \omega)\} dS_c dS_c, \tag{6.13}$$

where, considering the volume velocity $U_r(\boldsymbol{x}, \omega) = \dot{u}_r(\boldsymbol{x}, \omega) dS_c$ produced by the radial vibration of the infinitesimal area of the cylinder $\mathrm{d}S_c$, the acoustic impedance $Z(\boldsymbol{x}|\boldsymbol{x_a}, \omega)$ is given by [6,55]:

$$Z(\boldsymbol{x_a}|\boldsymbol{x_c}, \omega) = \frac{p(\boldsymbol{x_a}, \omega)}{U_r(\boldsymbol{x_c}, \omega)} \tag{6.14}$$

and applying Kirchhoff-Helmholtz, the acoustic impedance for the cylinder results given by:

$$Z(\boldsymbol{x_a}|\mathbf{x_c}, \omega) = -i\frac{\omega\rho}{2\pi^2} \sum_{n=0}^{N'} \varepsilon_m \cos\left(n\left(\theta_i' - \theta\right)\right) \overline{I_n}\left(z_i' - z_j\right). \tag{6.15}$$

Here, Equation (6.6) is employed using the approximation derived in (6.8). The integral expression for the time averaged total sound power radiation can be approximated into a Riemann sum over the mesh of $N = 62 \times 20 = 1240$ too, such that the time averaged total sound power radiation is given by:

$$\begin{aligned} \bar{P}(\omega) &= \frac{1}{2}\frac{S_c}{N} \sum_{j=1}^{N} \mathcal{R}e\left\{p_j^*(\omega)\dot{u}_{rj}(\omega)\right\} \\ &= \frac{1}{2}\frac{S_c}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathcal{R}e\left\{\dot{u}_{ri}^*(\omega) Z_{ij}^*(\omega) \dot{u}_{rj}(\omega)\right\}, \end{aligned} \tag{6.16}$$

where $p_\mathrm{j}(\omega)$ and $\dot{w}_\mathrm{j}(\omega)$ are the complex amplitudes of the sound pressure and radial velocity at the centre of the $j-th$ element and:

$$Z_{\mathrm{ij}}(\omega) = \frac{p_{\mathrm{i}}(\omega)}{\dot{u}_{\mathrm{rj}}(\omega)} = -i\frac{\omega\rho_0 S_c}{2N\pi^2}\sum_{n=0}^{N'} \varepsilon_m \cos\left(n\left(\theta_i' - \theta_j\right)\right) \overline{I_n}\left(z_i' - z_j\right). \quad (6.17)$$

Again, the complex velocities at the centre points of the mesh of elements can be obtained from the dynamic flexibility relation (5.30) such that:

$$\bar{P}(\omega) = \frac{1}{2}\frac{S_c}{N}\omega^2 \sum_{\mathrm{i}=1}^{\mathrm{N}}\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathcal{R}e\left\{f_r^*(\omega)\alpha_{\mathrm{i}}^*(\omega)Z_{\mathrm{ij}}^*(\omega)\alpha_{\mathrm{j}}(\omega)f_r(\omega)\right\}. \quad (6.18)$$

These summations are implemented with algebraic expressions based on the vectors of the complex velocities $\dot{\mathbf{u}}_r(\omega_k) = \lfloor \dot{u}_{\mathrm{r}1}(\omega_k) \quad \cdots \quad \dot{u}_{rN}(\omega_k) \rfloor^{\mathrm{T}}$ and complex pressures $\boldsymbol{p}(\omega_k) = \lfloor p_1(\omega_k) \quad \cdots \quad p_N(\omega_k) \rfloor^{\mathrm{T}}$ at the grid of centres points and $\boldsymbol{\alpha}(\omega_k) = \lfloor \alpha_1(\omega_k) \quad \cdots \quad \alpha_N(\omega_k) \rfloor^{\mathrm{T}}$ is the vector with the dynamic flexibilities such that:

$$\bar{P}(\omega_k) \cong \frac{1}{2}\frac{S_c}{N}\mathcal{R}e\{\boldsymbol{p}^{\boldsymbol{H}}(\omega_k)\dot{\mathbf{u}}_r(\omega_k)\} = \frac{1}{2}\frac{S_c}{N}\mathcal{R}e\{\dot{\mathbf{u}}_r^H(\omega_k)\boldsymbol{Z}^H(\omega_k)\dot{\mathbf{u}}_r(\omega_k)\}, \quad (6.19)$$

where $(\quad)^{\mathrm{H}}$ is the Hermitian-transpose matrix operator and $\boldsymbol{Z}(\omega_k)$ is a square matrix with the acoustic impedances given in Equation (6.17). For the special case at hand, the $\mathbf{Z}(\omega_k)$ matrix is symmetric and thus the time average total sound power radiation can be expressed in the following compact matrix form [58]:

$$\bar{P}(\omega_k) \cong \dot{\mathbf{u}}_r^H(\omega_k)\boldsymbol{R}^H(\omega_k)\dot{\mathbf{u}}_r(\omega_k) = \omega_k^2\boldsymbol{\alpha}^H(\omega_k)\boldsymbol{R}^H(\omega_k)\boldsymbol{\alpha}(\omega_k)f_r^2, \quad (6.20)$$

where $\boldsymbol{R}^H(\omega_k)$ is the so called radiation matrix, whose elements are defined as follows:

$$R_{ij}(\omega) = \frac{1}{2}\frac{S_c}{N}\mathcal{R}e\left\{Z_{\mathrm{ij}}(\omega)\right\}. \quad (6.21)$$

For completeness, the sound power radiation is calculated from both the camera vibration measurements and the laser vibrometer measurements. In the latter case, Equation (6.20) is still used recalling that the vector of admittances can be retrieved straightforwardly from that of the mobilities measured with the laser: $\boldsymbol{\alpha}(\omega_k) = \mathbf{Y}(\omega_k)/(i\omega_k)$.

As shown in Figure 40, the total sound power radiation has been measured in an anechoic room with a microphone array that has generated a total of 48 *FRFs* given by the ratios of the complex amplitudes of the sound pressures, $p_i(\omega_k)$, measured at the centres of the 6×8 quadrilateral elements the radiation surface has been divided into and the complex amplitude of the radial force exerted on the cylinder, $f_r(\omega_k)$:

$$H_i(\omega_k) = \frac{p_i(\omega_k)}{f_r(\omega_k)}. \tag{6.22}$$

Hence, the time-averaged total sound power radiation has been derived assuming plane waves and using Equation (4.46) and Equation (4.50). The integration over the radiating surface Ω_a has been substituted into a Riemann mid-point integration over the mesh of elements N_a the surface has been divided into obtaining:

$$\bar{P}(\omega_k) \cong \frac{1\Omega_a}{2\mathrm{N_a}\rho c}\sum_{\mathrm{i=1}}^{\mathrm{N}_a}|H_i(\omega_k)|^2|f_r(\omega_k)|^2. \tag{6.23}$$

Moving to the results, the sound power levels derived using Equation (4.53) from camera, laser and microphones measurements, the latter one taken accordingly to the grid of points described in Section 6.1, are now presented. As anticipated above, the derivation of the sound power level from camera and laser data has been done using the grid of points on the surface of the cylinder S_c, as depicted in Figure 47a. The orange line in Figure 48, shows the narrow-band spectrum based on the camera measurements whereas the blue line in Figure 49 shows the narrow-band spectrum based on the laser vibrometer measurements. The violet line in both Figures show the narrow-band spectrum derived from the measurements in the anechoic room taken with the array of microphones using Equation (6.23).

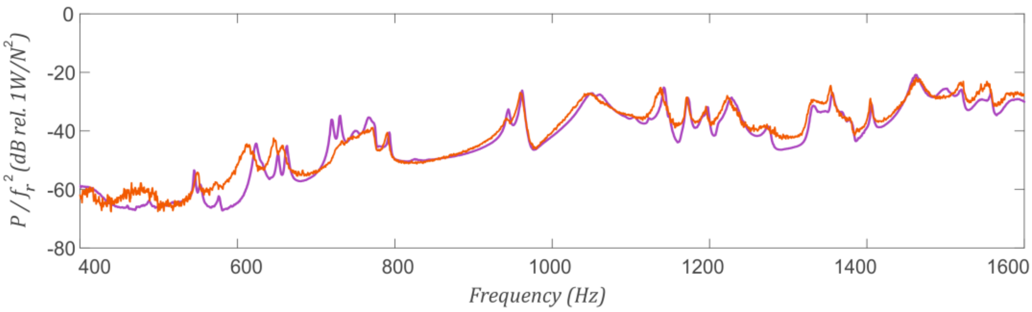


Figure 48: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and calculated from measurements with camera (orange line). Case with collapsed measurement surface

In general, the spectra obtained from the vibration measurements agree well with those obtained from the acoustic measurements. The spectrum produced from the camera measurements (orange line) shows only minor deviations at frequencies below 700 Hz compared to that derived from the acoustic measurements in the anechoic room (violet line). Also, it does not show deep throughs as those visible in the spectrum calculated from the acoustic measurements. These differences are in any case rather small and occur for sound power radiation values, that, as will be discussed below, have little influence on the third-octave values of the spectrum. The spectrum obtained from the laser vibrometer measurements (blue line) also agree well with that obtained from the acoustic measurements in the anechoic room (violet line). It shows slightly sharper resonance peaks, but, again, the differences from the spectrum measured with the microphones are small. In this case, there is a little mismatch between the spectra obtained from the laser measurements and from the acoustic measurements in the frequency band comprised between 640 and $680\ Hz$. Nevertheless, considering the three spectra in third octave bands, which are depicted in Figure 50 (orange bars from camera measurements, blue bars from laser measurements, violet bars from microphone measurements), except for the third octave band at $630\ Hz$, the differences between the spectral bars obtained from the vibration measurements and those obtained from the acoustic measurements are contained within a 1 to $2\ dB$.

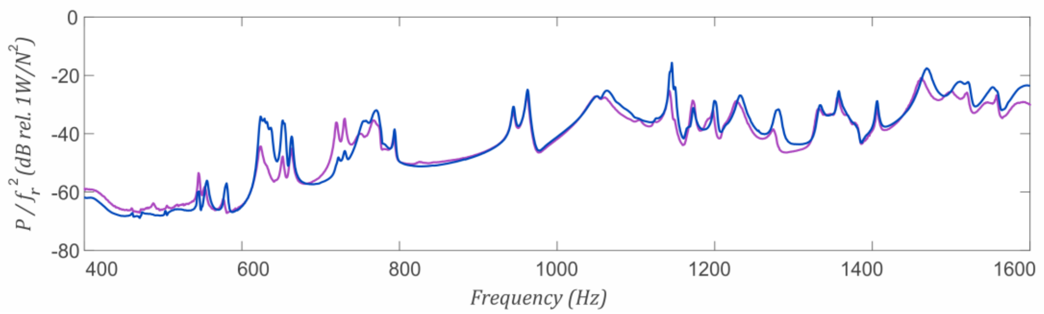


Figure 49: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and calculated from measurements with laser vibrometer (blue line).

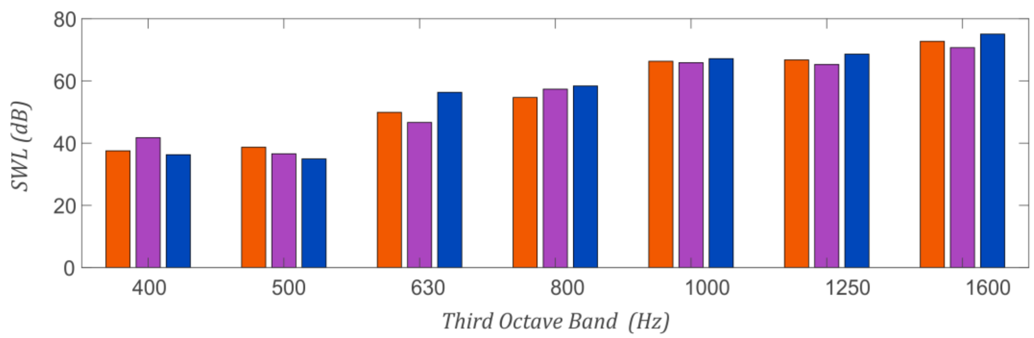


Figure 50: Third-octave band spectrum of the total sound power radiation from measurements with microphones (violet bars) and calculated from measurements with laser vibrometer (blue bars) and camera (orange bars) – $dB\ ref.\,10^{-12}\ W$.

Nevertheless, it should be highlighted that the *acoustic-based* measurement also offers an estimate, which, as a matter of fact, is based on a relatively small number of measurement points (the 6×8 measurement points depicted in Figure 47b) and relies on the quality of the anechoic room, which may be affected by poor absorption in specific frequency bands as well as flanking noise effects (outdoor and indoor, including the equipment used for the measurement). On the contrary, the proposed measurement approach relies on a rather dense grid of 62×20 vibration measurements, which in general is not affected by background acoustic noise effects, unless they are so high to alter the vibration of the structure. Again, in general, the camera vibration measurements are affected by pixelization errors, which principally depend on the time and spatial resolution of the camera. In this study a rather fast camera was employed, which guaranteed a correct reconstruction of the vibration field up to $2\ kHz$. The camera was also characterised by a rather high spatial resolution, which ensured an accurate reconstruction of the vibration amplitude. In this respect it is important to highlight, as done for the sound field measurements, that the errors arising from the pixelization of the displacements should have a random nature and thus, when the Kirchhoff-Helmholtz integral formulation is implemented to derive the total sound power radiation, these errors tend to cancel out such that, compared to that of the vibration measurement, the accuracy of the sound radiation prediction should be higher.

For completeness, beside the proposed approach based on the integration of the sound intensity over the radiating surface S_c (see Figure 47a), the total sound power radiation has also been derived with the classical approach, which depends on the integration of the sound intensity over a large surface Ω_a that encloses the radiating object, that is a large cylinder in this case (see Figure 47b). Thus, recalling Equation (4.45) and considering the fact that, in the far-field, the sound waves locally resemble a plane wave over the measurement surface Ω_a such that the particle velocity has radial direction given by Equation (4.46), the instantaneous total sound power radiation considering time-harmonic sound waves results expressed by Equation (4.50). In principle, the measurement surface should enclose the flexible cylinder and the rigid baffles, thus it should be a cylinder with large diameter D_a and infinite length. In this study, a cylinder with finite length H_r has been considered instead, which, nevertheless, covers the whole flexible cylinder and a portion of the baffles sufficient to detect the sound field radiated by the structure and scattered by the baffles. As done above for the integration over the sound radiation surface S_c in Equation (6.13), the total sound power radiation can be expressed in terms of the following double integration:

$$\bar{P}(\omega) = \frac{1}{2} \int_{\Omega_a} \int_{S_c} \mathcal{R}e\{\dot{u}^*{}_r(\boldsymbol{x},\omega) Z^*(\mathbf{x_a}|\mathbf{x},\omega) \dot{u}_r(\mathbf{x},\omega)\} \mathrm{d}S_c d\Omega_a. \tag{6.24}$$

As highlighted in the formula, in this case Equation (6.23) provides an approximated result since a finite radiation surface Ω_a is considered rather than an infinitely long cylinder. As shown in Figure 47b, to simplify the calculus, the integrals over the measurement surface Ω_a and the radiation surface S_c have been converted into Riemann midpoint sums across the meshes of $N_a = 6 \times 8 = 48$ and $N = 62 \times 20 = 1240$ elements, such that:

$$\bar{P}(\omega) \cong \frac{\Omega_a}{2N_a \rho c} \sum_{i=1}^{N_a} |p_i(\omega_k)|^2$$
$$= \frac{\Omega_a}{2N_a \rho_0 c} \sum_{\mathrm{i}=1}^{\mathrm{N}} \sum_{\mathrm{j}=1}^{\mathrm{N}_a} \dot{w}_{\mathrm{i}}^*(\omega) Z_{\mathrm{ij}}^*(\omega) Z_{\mathrm{ij}}(\omega) \dot{w}_{\mathrm{i}}(\omega). \tag{6.25}$$

Another time, the complex velocities at the centre points of the 62×20 mesh of elements in the cylinder surface can be obtained from the dynamic flexibilities in (5.30) derived from the camera measurements.

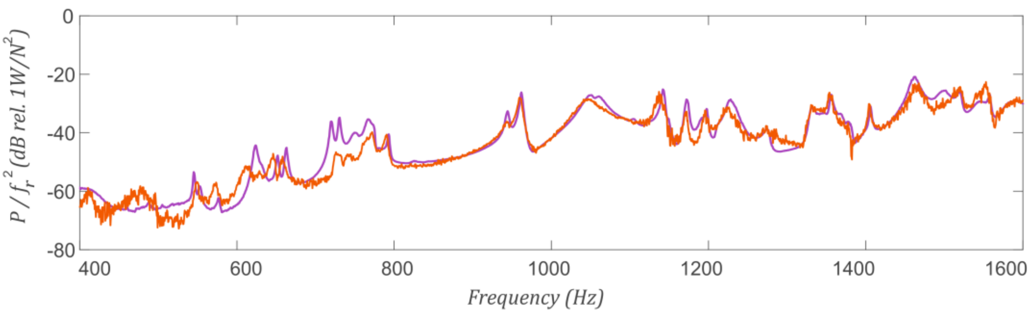


Figure 51: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and reconstructed at microphones positions from measurements with camera (orange line).

Moving another time to the results, the sound power levels derived using Equation (4.53) from camera, laser and microphones measurements are now presented. In this case, the derivation of the sound power level from camera, laser and microphones data has been done using the mesh of 6×8 points depicted in Figure 47b using the classical formulation that integrates the sound radiation over a far-field measurement surface Ω_a. Again, the orange line in Figure 51 shows the narrow-band spectrum based on the camera measurements whereas the blue line in Figure 52 shows the narrow-band spectrum based on the laser vibrometer measurements. Also in this case, the violet lines in both figures show the narrow-band spectrum derived from the measurements in the anechoic room taken with the array of microphones using Equation (6.23). In this case too, the spectra obtained from the vibration measurements agree well with those obtained from the acoustic measurements. There are still some critical issues at low frequencies, below about $800\,Hz$. Overall, the spectra obtained from vibration measurements in Figures 51 and 52 agree less well with that obtained from the acoustic measurement than the spectra obtained from vibration measurements in Figures 48 and 49. This suggests that the derivation of the total sound power radiation from the integral of the sound intensity over the surface of the cylinder S_c, using Equation (6.20), is more accurate than the classical approach based on the integration of the sound radiation over a far-field measurement surface Ω_a. This is mainly due to the fact that in the former case the integration has been approximated by a Riemann sum over a smaller surface ($S_c \ll \Omega_a$) with a much larger number of points ($N \gg N_a$).

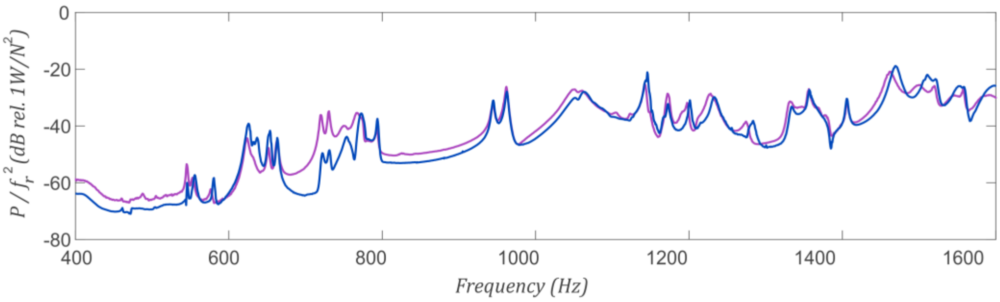


Figure 52: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and reconstructed at microphones positions from measurements with laser vibrometer (blue line).

For the sake of completeness, the one-third octave band spectrum is herein included and presented in Figure 53, where the orange bars represented the sound power level from camera measurements, the blue bars from laser measurements and the violet bars from microphone measurements.

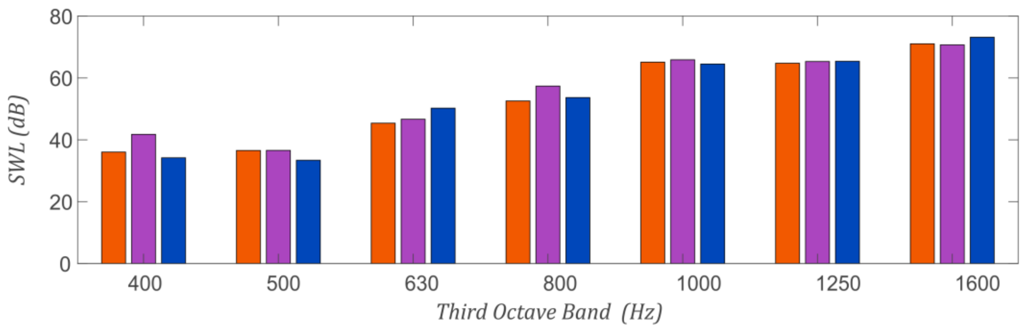


Figure 53: third-octave band spectrum of the total sound power radiation from measurements with microphones (violet bars) and reconstructed at microphones positions from measurements with laser vibrometer (blue bars) and camera (orange bars) – $dB\ rel.\,to\ 10^{-12}\ W$.

SUMMARY OF THE CHAPTER

This chapter has presented an experimental and analytical investigation of the sound radiation from a baffled cylindrical structure. Dedicated test rigs were introduced to measure both the acoustic field and the radiated sound power, providing a controlled framework for validating the proposed methodologies.

The Kirchhoff–Helmholtz formulation was employed to derive the acoustic pressure field from the measured structural vibration, and the resulting sound radiation maps showed good agreement with those obtained from direct microphone measurements. In addition, sound power level spectra were derived using a surface-based formulation, demonstrating that vibration-driven approaches can accurately reproduce both narrow-band and third-octave representations of the radiated sound power.

Overall, the results obtained for this analytically well-defined structure confirm the validity of the proposed optical and vibro-acoustic framework. They provide a reference validation case and establish the basis for its extension to more complex and realistic structures, which is addressed in the subsequent chapter.

7
SMALL BOX SOUND RADIATION

This chapter focuses on the experimental and numerical investigation of the acoustic radiation from a vibrating small box structure. Unlike the previous chapter focused on the sound radiation of a cylindrical shell, the box represents a more complex engineering component with structural discontinuities, sharp edges, and non-ideal boundary conditions, providing an opportunity to test the robustness and applicability of the methodology beyond idealized laboratory geometries. The discussion begins with the description of the measurement setup, where an array of microphones was arranged in a free-field configuration according to *ISO 3744:2010* standards [54,105], providing a reference dataset for the evaluation of radiated sound pressure. Complementary vibration measurements were performed using both a scanning laser vibrometer and a high-speed camera, with the aim of reconstructing the radiated field through the boundary element method (*BEM*). The implementation of the *BEM* algorithm is first described, including the discretization with a triangular mesh and the treatment of edge conditions. Particular attention is devoted to the validation of the approach, starting from simplified simulations that quantify the effect of constraining edge velocities on the accuracy of the reconstructed acoustic response. Subsequently, the reconstructed sound pressure fields are compared with direct microphone measurements at selected control points, allowing a detailed assessment of the accuracy and limitations of both the laser- and camera-based approaches across a wide frequency range. The analysis then extends to the derivation of the radiated sound power. Two methodologies are presented and critically compared: the classical *ISO-based* approach, in which the sound power is obtained by integrating the sound intensity over a far-field measurement surface, and a novel approach for the reconstruction of the sound field with the *BEM.* Part of this novel approach relies on the assumption that the structural velocity coincides with the acoustic particle velocity at the boundary. [106]. The chapter concludes with a comparative discussion of the results obtained in both narrow-band and third-octave representations. This provides a comprehensive evaluation of the strengths and limitations of each approach in estimating sound power radiation from vibrating structures, even in the presence of structural discontinuities.

7.1 TEST RIG: MICROPHONE ARRAYS

The total sound radiation was measured with the vertical array of microphones shown in Figure 54b. The measurement was carried out according to the *ISO 3744:2010* guidelines [54,105] for the estimation of sound radiation in an anechoic room. In general, this measurement is carried out in a semi-anechoic room by placing the test object on the floor and taking sound pressure measurements on a grid of points lying on a large hemispherical surface where the sound field produced by the radiator can be locally approximated as a plane wave such that the radiated sound power can be derived from measurements of the sound pressure only [10]. Again, given the peculiar shape of the small box radiator considered here (which is taller than it is wide), as can be seen in Figure 54b, the measurement was carried out in a semi-anechoic room and the sound pressures were taken on a large, closed cylindrical surface (see Figure 54a), which was extended over the box. In this way, the direct sound radiation by the box was measured appropriately. For the given dimensions of the box and of the whole anechoic room, following *ISO 3744:2010* standards, the measurement cylindrical surface was taken to have a diameter $D_s = 2(l_1/2 + d_1) = 2(240/2 + 1000) = 2240\ mm$ and a height $H_s = 1420\ mm$. The sound pressures were measured in a $12 \times 8 = 96$ grid of points (8 in circumferential, 8 in axial and 4 in radial direction). The measurement points were located enforcing the following standard conditions: $n_s \geq H_s/0.5 \geq 4$ and $n_T \geq n_s/2$ where n_s denotes the number of measurement points along the vertical side of the cylinder, n_T is the number of measurement points on the top surface. Therefore, the measurements points were located at the centre of a $8 \times 8 = 64$ regular mesh of quadrilateral elements on the vertical surface and at the centre of a $4 \times 8 = 32$ mesh of quadrilateral/triangular elements on the horizontal top surface.

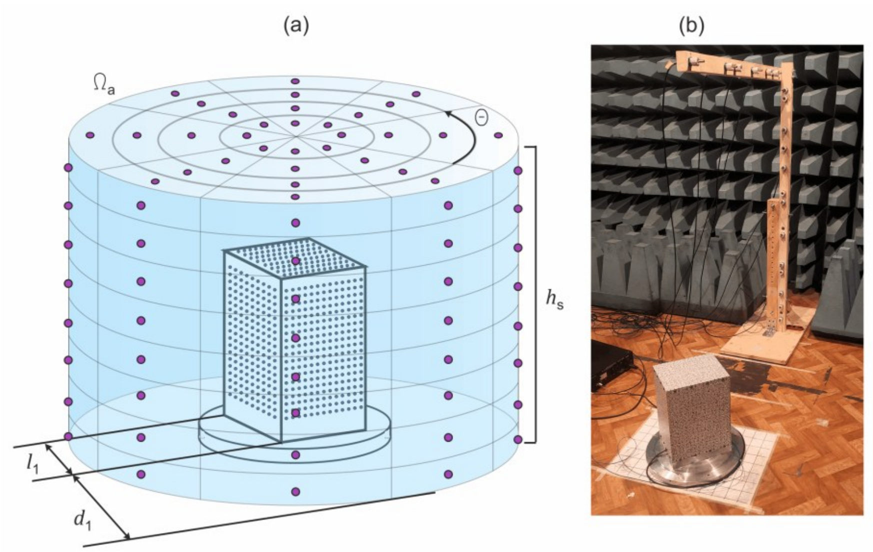


Figure 54: Setup for acoustic measurements. a) scketch of all ISO 3744:2010 standard measurement points, b) Array of 12 microphones in a semi-anechoic room.

As can be notice in Figure 54b, the measurements were taken with the vertical array of 8 microphones and the radial array of 4 microphones. As done with the laser vibrometer and the camera measurements, here too, the array of microphones was fixed and the box was rotated to acquire the circumferential positions. More specifically, the acquisition from the array of microphones was repeated 8 times with $45\ deg$ rotations of the small box. The measurements were carried out over a $200\ -\ 2000\ Hz$ frequency range such that the acoustic wavelength spanned the range between $1.72\ m$ and $0.17\ m$.

7.2 SOUND PRESSURE DERIVATION

As described in Chapter 5, the mobility Frequency Response Functions (*FRFs*) of the normal velocities per unit force excitation Y_b were first acquired with laser measurements and reconstructed from camera measurements using a frequency-domain triangulation technique at the five grids of measurement points mentioned in Table 3. As anticipated in Chapter 5, these measurement points were located at the vertices of triangular elements, allowing the application of the boundary element method which is particularly suitable for geometries that do not permit an analytical solution (as described in Chapter 4) of the acoustic radiation problem. The Boundary Element Method enables the discretization and simplification of the Kirchhoff–Helmholtz integral equation, while the selection of points at the vertices of the triangles facilitates numerical convergence during the iterative computational procedure. The *BEM* employed in this study [107] implements a direct solution approach for acoustic radiation, as presented in Section 4.4.

MESH IMPLEMENTATION

The implementation of a direct collocational *BEM* requires the radiating object to be represented by a closed mesh made with quadrangular or triangular elements, on which specific boundary conditions are imposed at the nodes. To build up the mesh, *Gmsh* software was employed [108]. *Gmsh* is an open-source three-dimensional finite element mesh generator that integrates a *CAD engine* and *post-processing* functionalities. It has been developed with the objective of offering a fast, lightweight, and user-friendly tool for mesh generation, featuring parametric input capabilities and versatile visualization options. The software is organized into four main modules, geometry, mesh, solver, and post-processing, which can be operated via the graphical user interface, through command-line instructions, by means of text files in its native scripting format (*.geo* files), or programmatically through application programming interfaces available in *C++*, *C*, *Python*, *Julia*, and *Fortran*.

As shown in Figure 55, the mesh construction process followed a structured meshing approach on the five faces of the box to ensure precise alignment between the measurement points and the corresponding nodes on the $3D$ *CAD* model. The basement was also modelled to account for scattering effects, but in this case, a structured mesh configuration was not maintained. To achieve the required computational accuracy, the element density criterion [88,109] was strictly enforced, ensuring that the mesh contained at least six elements per half-wavelength at the highest frequency of interest.

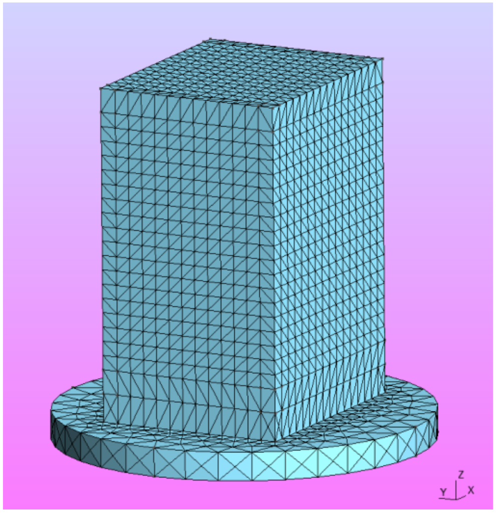


Figure 55: semi-structured mesh of the box made with gmsh software.

This criterion is essential to accurately capture the spatial variations of the acoustic field and to prevent numerical dispersion errors. In addition, triangular elements were chosen to improve the convergence of the solution [110].

BOUNDARY ELEMENT IMPLEMENTATION

Starting from the Kirchhoff-Helmholtz equation (4.29) the solution for acoustic radiation was derived using a direct collocational *BEM* approach. The boundary conditions used to solve the wave equation in this case are the Sommerfeld radiation condition at infinity, and the constraint imposed on the normal component of the velocity:

$$\begin{cases} \lim_{\boldsymbol{r}\to\infty} |\boldsymbol{r}| \left(\frac{\partial p}{\partial \boldsymbol{r}} + ikp\right) = 0 \\ \\ \dot{u}_r(\omega) = \frac{i}{\rho\omega}\frac{\partial p}{\partial \boldsymbol{n}} \end{cases} \tag{7.1}$$

Moreover, considering that the structure is on a rigid reflective floor, the Green's function in Equation (4.29) is set to be the half-space Green's function $G(\mathbf{x}_\mathrm{n}, \mathbf{x}_\mathrm{m}) = \frac{e^{-jk|x-x_a|}}{4\pi|x-x_a|} + \frac{e^{-jk|x'-x_a|}}{4\pi|x'-x_a|}$ where $\boldsymbol{x}'$ represents the mirror source location. These boundary conditions allow the derivation of the acoustic pressure $p_m(\boldsymbol{x_a}, \omega)$ at any field point within the unbounded domain, which is directly related to the pressure distribution $p_b(\boldsymbol{x}, \omega)$ and the normal velocity distribution $\dot{u}_r(\boldsymbol{x}, \omega)$ on the closed surface of the radiating object. For this reason, solving the integral to obtain $p_m(\boldsymbol{x_a}, \omega)$ requires a *two-step* procedure (as discussed in Chapter 4), which is typically adopted in the Boundary Element Method. In the first step, an approximation of the boundary variables on the surface is determined. This approximation is based on the expansion of both the boundary surface geometry and the boundary variables in terms of a set of prescribed linear shape functions (Equation (4.32)), locally defined within small subsurfaces, referred to as elements e, of the boundary surface. By discretizing the surface into elements, n nodes are also introduced at the vertices of each element. In this study, for each boundary variable, one degree of freedom is assigned per node. Accordingly, linear shape functions have been adopted which for the reference element have a form as:

$$\boldsymbol{\phi}(\xi,\eta) = \begin{bmatrix} \phi_1(\xi,\eta) \\ \phi_2(\xi,\eta) \\ \phi_3(\xi,\eta) \end{bmatrix} = \begin{bmatrix} 1-\xi-\eta \\ \xi \\ \eta \end{bmatrix}. \tag{7.2}$$

In this way, the original problem of determining the variables on the continuous boundary surface is transformed into an equivalent problem of determining the variables at discrete positions on the surface (see Ref [58]). This results in a set of algebraic equations forming the direct boundary element model (4.39), which, for the present case study, is expressed as:

$$[\boldsymbol{A}]\boldsymbol{p_b} = i\rho\omega[\boldsymbol{B}]\dot{\boldsymbol{u}}_{\boldsymbol{r}}, \tag{7.3}$$

where $[\boldsymbol{A}]$ and $[\boldsymbol{B}]$ are the matrices introduced in Section 4.4.

In the second step, the acoustic pressure at the selected points $p_m(\boldsymbol{x_a}, \omega)$ in the unbounded fluid domain is computed from the boundary integral formulation, using the boundary surface results obtained in the first step:

$$\boldsymbol{p_m}(\boldsymbol{x_a}, \omega) = [\boldsymbol{A_m}]\boldsymbol{p_b}(\boldsymbol{x}, \omega) + ik\rho c\omega[\boldsymbol{B_m}]\dot{\boldsymbol{u}}_{\boldsymbol{r}}(\boldsymbol{x}, \omega), \tag{7.4}$$

where $[\boldsymbol{A_m}]$ and $[\boldsymbol{B_m}]$ represent the corresponding matrices derived from the first-step results.

Since the laser vibrometer and high-speed camera measurements were performed on points located within the faces of the box, without providing results for the edges, an initial simulation was carried out to assess the effect of the edges on the acoustic radiation derived through the *BEM*. To this end, the same box tested in the laboratory was modelled in the *Siemens NX* environment, and a structural simulation was performed in which the box was constrained only at its base. The resulting nodal normal velocities were then used within the *BEM* algorithm to create two case studies: the first utilised the complete set of nodal velocity data obtained from the *NX* simulation, while in the second, the normal velocities at the edge nodes were set to zero. These preliminary tests enabled a direct comparison between the two cases, aimed at evaluating the impact of assuming zero nodal normal velocities along the edges of the box. The comparison showed that, for the case under investigation, imposing zero normal velocities on the edges did not significantly affect the overall results in the derivation of the acoustic pressure at the 96 points in the free-field domain. The model, designed to closely resemble the laboratory box, was analysed primarily at resonance frequencies exhibiting an odd number of lobes in the deformation patterns of the individual faces. This configuration represents the most critical case, as the radiation values do not cancel out when the number of radiators is odd. An example of seven tested frequencies is presented in Figure 56, where the orange dots represent the *SWL* derived using all the nodes information obtained from *NX*, while the blue dots represent the *SWL* derived forcing the edge velocities to be equal to zero. Nevertheless, this figure shows that the deviation from the sound power level between the two test cases remains negligible.

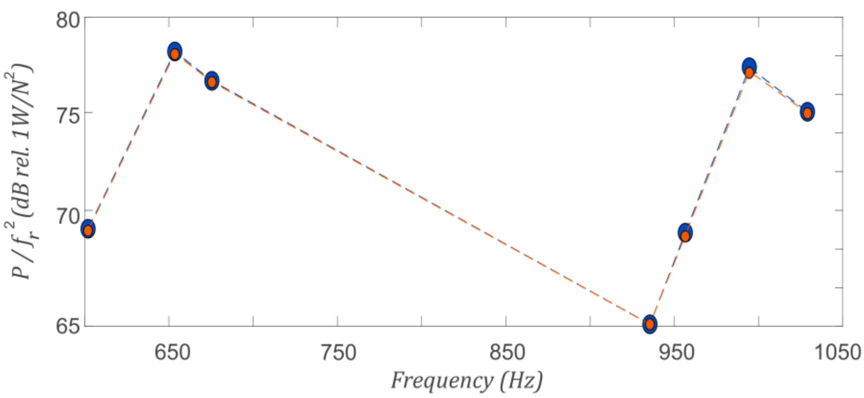


Figure 56: comparison between SWL derived from NX simulations. Orange dots using all nodal data, blue dots forcing edges to zero velocity.

Moving now to the results on the box structure, Figure 57 illustrates the subset of control points selected from the 96 available measurement positions defined according to the *ISO 3744:2010 standard*, which have been used for a detailed comparison of the sound radiation data. The chosen control points are located at representative positions on the box surfaces, enabling the assessment of the agreement between the directly measured and the reconstructed acoustic fields (using *BEM* algorithms) under different geometric and acoustic conditions. The sound radiation is assessed here at six control positions, expressed in terms of the time-averaged sound pressure per unit force excitation, either measured directly with a microphone (violet lines in Figures 58a–f and 59a–f) or derived from Equation (7.4) using the vibration *FRFs* acquired with the laser vibrometer (blue lines in Figures 58a–f) or reconstructed from the camera measurements (orange lines in Figures 59a–f).

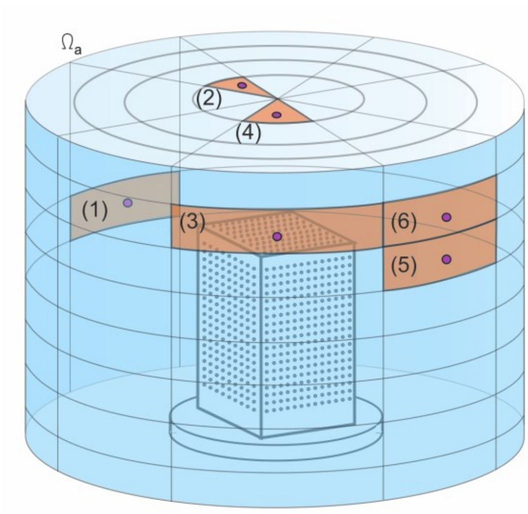


Figure 57: control points in the far-field domain for comparison between measured sound pressures and simulated using BEM.

For the two control points located on *face 2*, opposite to the excitation face, *point (1)* (Figures 58a and 59a) and *point (2)* (Figures 58b and 59b), the spectra derived from the laser vibrometer (blue line) show a good agreement with the microphone measurements (violet line), reproducing both the main resonances and the antiresonance troughs. Small mismatches occur in narrow frequency bands, possibly due to environmental noise, slight changes in the modal response during the measurement campaign, or the influence of flanking transmission paths. In general, the camera-derived spectra (orange line) overlap quite well with the microphone spectra, but show more pronounced deviations at higher frequencies, where the spatial resolution limitations of the camera and the noise present in the kinetic energy spectrum (discussed in Chapter 5) may affect the reconstruction accuracy. For the two control points located between *faces 3* and *4*, *point (3)* (Figures 58c and 59c) and *point (4)* (Figures 58d and 59d), the agreement between blue and violet curves remains consistent, with most resonance peaks well aligned in frequency and amplitude.

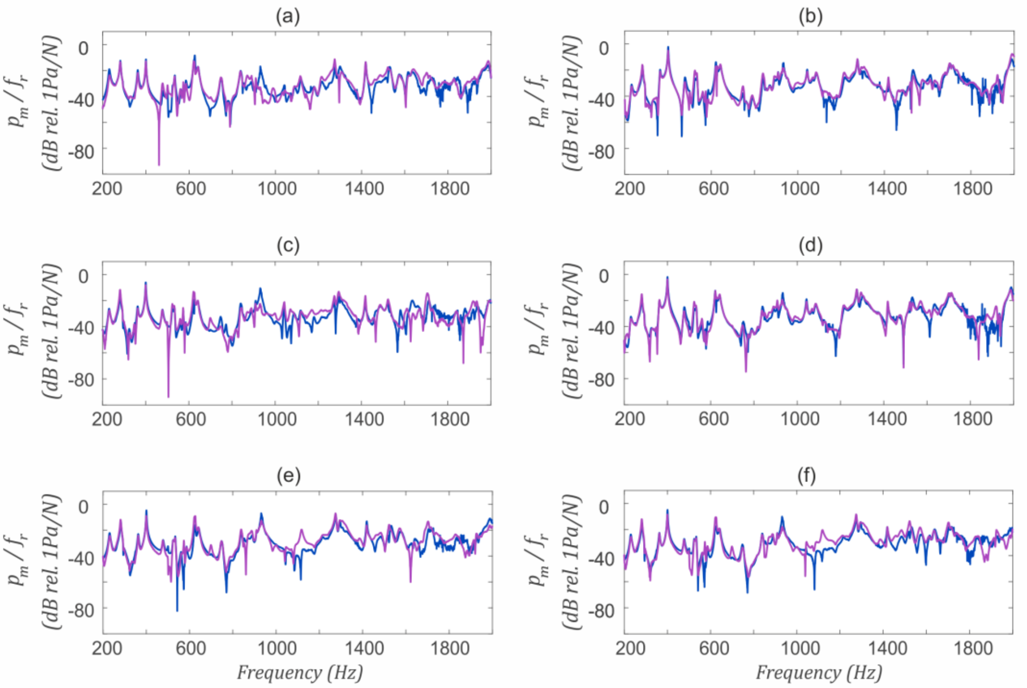


Figure 58: spectra of sound pressures at six monitor points measured with microphones (violet line) and derived from laser vibrometer measurements (blue line).

The orange curves, while still capturing the main spectral features, show a slightly reduced accuracy in matching peak amplitudes, especially above $1200\ Hz$. This behaviour could be linked to phase reconstruction errors, reduced measurement sensitivity in that frequency range, or even slight variations in the structural response between the acquisition of the different datasets. Finally, for the two control points on *face 4*, the excitation face, *point (5)* (Figures 58e and 59e) and *point (6)* (Figures 58f and 59f), both the laser vibrometer and the camera reconstructions, are able to reproduce the microphone spectra with reasonable fidelity. The blue lines provide the closest match in terms of both peak positions and amplitudes, whereas the orange lines occasionally underestimate the resonance amplitudes, particularly at mid–high frequencies. The higher spectral roughness visible in the orange curves is again partly attributable to the noise in the kinetic energy spectrum described in Chapter 5. Comparing the results from the laser vibrometer (blue) and the camera (orange) over all control points, both approaches reproduce the overall shape and dominant resonances of the microphone measurements. The *laser-based* method generally achieves a closer match across the entire frequency range, while the *camera-based* method still shows promising performance given its non-contact, wide-field nature, making it a viable alternative for estimating sound radiation over a broad frequency range with acceptable deviations.

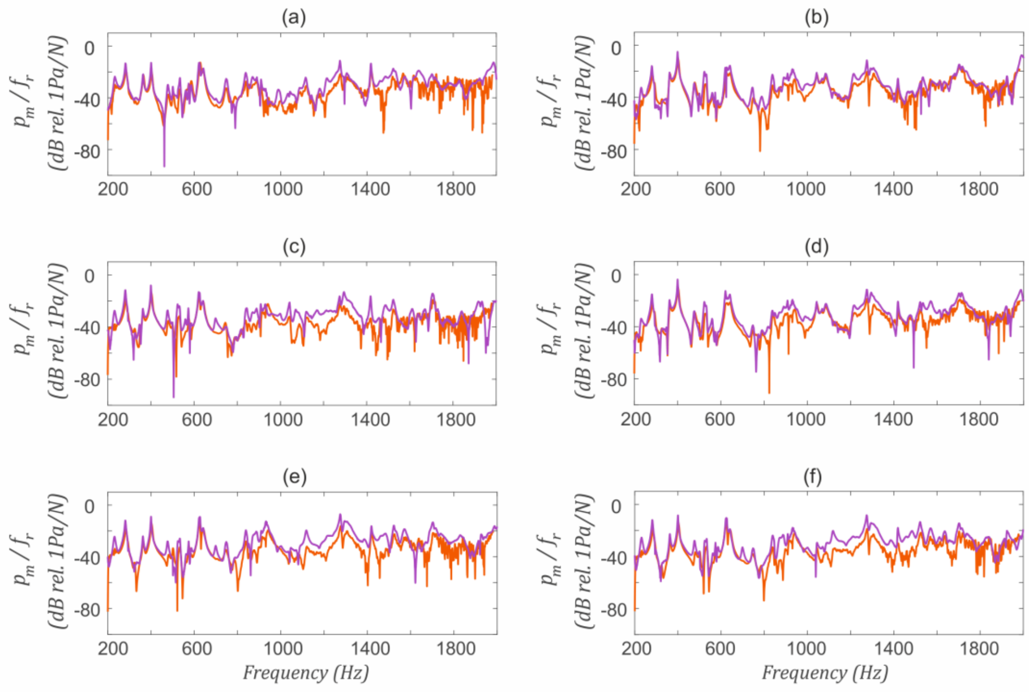


Figure 59: spectra of sound pressures at six monitor points measured with microphones (violet line) and derived from laser vibrometer measurements (blue line).

These results will be further exploited in the following sections to produce sound power level spectra for the tested configuration. These considerations are representative of the trends observed across the complete set of 96 measurement points, which are not individually reported here to avoid unnecessarily lengthening the presentation of results.

7.3 SOUND POWER LEVEL DERIVATION

As described in Chapter 4, and already implemented in Chapter 6 for the baffled cylinder (whose sound radiation can be derived analytically), the total averaged sound power radiated by a vibrating structure can be obtained from Equation (4.49), once the surface pressures of the box have been derived through the Kirchhoff–Helmholtz formulation discretized in Section 7.2 using the Boundary Element Method. In the present study, however, a novel and simpler approach is introduced, whereby, as illustrated in Figure 60, the closed surface Ω_a is collapsed onto the surface of the radiating structure itself, namely the surface of the flexible box S_b. In this configuration, the fluid particle velocity $v(\boldsymbol{x})$ over the radiating surface coincides with the normal (i.e., flexural) velocity $\dot{u}_r(\boldsymbol{x})$ of the box walls, so that the instantaneous total radiated sound power is expressed as:

$$\bar{P}(\omega) = \frac{1}{2}\int_{S_b} \mathcal{R}e\{p_s^*(\boldsymbol{x},\omega)\dot{u}_r(\boldsymbol{x},\omega)\}dS_b, \tag{7.5}$$

where $(\quad)^*$ is the complex-conjugate of the argument.

Considering the Kirchhoff–Helmholtz formulation implemented in the *BEM* algorithm (Section 7.2), the total radiated sound power can be discretized over each triangular element of the box mesh shown in Figure 55. By introducing normalized coordinates (ξ,η), which allow the mapping of each triangular element onto a reference element and enable the interpolation of boundary variables through the adopted linear shape functions ϕ, the expression can be written in the form of a double integral as:

$$\bar{P}(\omega) = \frac{1}{2}\sum_{e=1}^{N_e}\int_0^1\int_0^{1-\eta} \mathcal{R}e\{p_s^*(\boldsymbol{x},\omega)\dot{u}_r(\boldsymbol{x},\omega)\}\,\det(\boldsymbol{J})\,\mathrm{d}\xi d\eta, \tag{7.6}$$

where det $(\boldsymbol{J})$ represents the determinant of the Jacobian matrix.

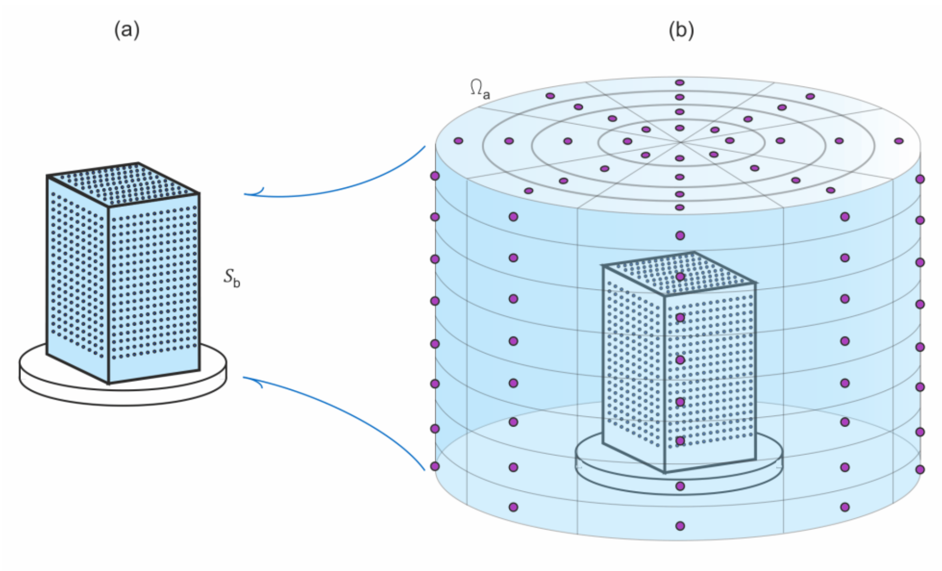


Figure 60: a) proposed near-field sound radiation measurement surface built by collapsed the classic far-field sound radiation measurement surface (b) on the box.

In accordance with the procedure described in Chapter 4, the value of a boundary variable at any position within a triangular element can be represented as a linear interpolation of its nodal values, weighted by the corresponding shape functions. By substituting Equation (4.32) into Equation (7.6), the following expression is obtained:

$$\bar{P}(\omega) = \frac{1}{2}\sum_{e=1}^{N_e}\int_0^1\int_0^{1-\eta} \mathcal{R}e\{\boldsymbol{p}_{\boldsymbol{se}}^{\boldsymbol{H}}(\omega)\boldsymbol{\phi}^{\boldsymbol{T}}(\xi,\eta)\boldsymbol{\phi}(\xi,\eta)\dot{\boldsymbol{u}}_{\boldsymbol{re}}(\omega)\}\det(\boldsymbol{J})\,\mathrm{d}\xi d\eta, \qquad (7.7)$$

where $(\ )^{\mathrm{H}}$ is the Hermitian-transpose matrix operator, $\boldsymbol{\phi}(\xi,\eta) = \lfloor \phi_1(\xi,\eta)\ \ \phi_2(\xi,\eta)\ \ \phi_3(\xi,\eta) \rfloor$ represents the vector of shape functions, $\dot{\mathbf{u}}_{\boldsymbol{re}}(\omega)$ is the vector of the complex velocities and $\boldsymbol{p}_{\boldsymbol{se}}(\omega)$ is the vector of complex pressures at the nodes of each element. It is recalled that, based on the experimental measurements obtained with both the laser vibrometer and the high-speed camera, the nodal velocities are related to the applied force through the vector of dynamic flexibilities $\boldsymbol{\alpha}_e(\omega)$, giving: $\dot{\mathrm{u}}_r(\omega) = i\omega\alpha(\omega)f_r$ as done for the baffled cylinder in previous chapters. However, for the sake of simplicity in the present derivation, the nodal velocity will be denoted simply as $\dot{\mathbf{u}}_r(\omega)$.

Equation (7.7) can be rewritten in the following form:

$$\begin{aligned}\bar{P}(\omega) &= \frac{1}{2}\sum_{e=1}^{N_e}\mathcal{R}e\left\{\boldsymbol{p}_{\boldsymbol{se}}^{\boldsymbol{H}}(\omega)\int_0^1\int_0^{1-\eta}\boldsymbol{\phi}^{\boldsymbol{T}}(\xi,\eta)\boldsymbol{\phi}(\xi,\eta)\det(\boldsymbol{J})\mathrm{d}\xi d\eta\,\dot{\boldsymbol{u}}_{\boldsymbol{re}}(\omega)\right\} \\ &= \frac{1}{2}\sum_{e=1}^{N_e}\mathcal{R}e\{\boldsymbol{p}_{\boldsymbol{se}}^{\boldsymbol{H}}(\omega)[\boldsymbol{C}_{\boldsymbol{e}}]\dot{\boldsymbol{u}}_{\boldsymbol{re}}(\omega)\}\end{aligned} \qquad (7.8)$$

Where the matrix $[\boldsymbol{C}_{\boldsymbol{e}}]$ has been introduced. Therefore, considering the summation over all the elements, the total sound power radiated can be written as:

$$\bar{P}(\omega) = \frac{1}{2}\mathcal{R}e\{\boldsymbol{p}_{\boldsymbol{s}}^{\boldsymbol{H}}(\omega)[\boldsymbol{C}]\dot{\boldsymbol{u}}_{\boldsymbol{r}}(\omega)\} \qquad (7.9)$$

In which $\boldsymbol{p}_{\boldsymbol{s}}(\omega) = \lfloor p_1(\omega)\ \ \cdots\ \ p_N(\omega)\rfloor^{\mathrm{T}}$ and $\dot{\mathbf{u}}_r(\omega) = \lfloor \dot{u}_{\mathrm{r}1}(\omega)\ \ \cdots\ \ \dot{u}_{rN_e}(\omega)\rfloor^{\mathrm{T}}$ represents the vectors of nodal pressures and nodal velocities respectively and $[\boldsymbol{C}]$ is the global matrix with the computed double integral for each element of the mesh. The global matrix $[\boldsymbol{C}]$ is assembled summing the elemental matrices in the relevant rows and columns. The use of normalized coordinates (ξ,η) and Gaussian quadrature for the numerical evaluation of the integrals ensures a consistent and accurate discretization, fully aligned with the approximation framework introduced in Chapter 4. Moreover, the $[\boldsymbol{C}]$ matrix, calculated for the box structure analysed in this thesis as the sparsity pattern reported in Figure 61.

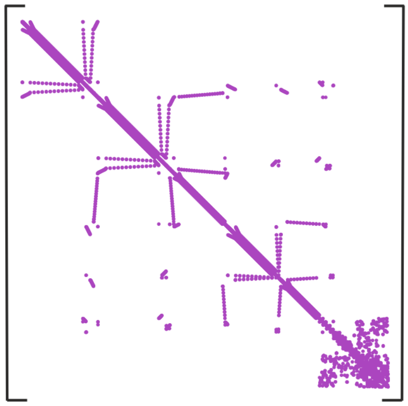


Figure 61: sparsity pattern of matrix C. Non-zero values are coloured while zero values are white.

Recalling Equation (7.3), the vector of nodal pressures can be derived from the vector of nodal velocities as follows:

$$\boldsymbol{p_s}(\omega) = i\omega\rho[\boldsymbol{A}]^{-1}[\boldsymbol{B}]\dot{\mathbf{u}}_r(\omega) = [\boldsymbol{Z}(\omega)]\dot{\mathbf{u}}_r(\omega), \tag{7.10}$$

where $[\boldsymbol{A}]^{-1}$ represents the inverse of the matrix $[\boldsymbol{A}]$ and $[\boldsymbol{Z}(\omega)]$ is the matrix of acoustic impedance *FRFs* as described in Section 6.3. Thus, Equation (7.9) becomes:

$$\bar{P}(\omega) = \frac{1}{2}\mathcal{R}e\{\dot{\boldsymbol{u}}_{\boldsymbol{r}}^{\boldsymbol{H}}(\omega)[\boldsymbol{Z}](\omega)[\boldsymbol{C}]\dot{\boldsymbol{u}}_{\boldsymbol{r}}(\omega)\}. \tag{7.11}$$

It is important to underline that, if the matrix $[\boldsymbol{C}]$ is symmetric (this could be achieved for example with a Galerkin approach usually implemented in indirect *BEM* algorithms) and then $[\boldsymbol{Z}(\omega)][\boldsymbol{C}]$ is symmetric too, the total sound power radiated can be rewritten in the following form:

$$\bar{P}(\omega) = \frac{1}{2}\mathcal{R}e\{\dot{\boldsymbol{u}}_{\boldsymbol{r}}^{\boldsymbol{H}}(\omega)[\boldsymbol{R}(\omega)]\dot{\boldsymbol{u}}_{\boldsymbol{r}}(\omega)\}, \tag{7.12}$$

where $\mathcal{R}e\{[\boldsymbol{R}(\omega)]\}$ is the classical sound radiation matrix.

As shown in Figure 54b, the total sound power radiation has been measured in a semi-anechoic room with a microphone array that has generated a total of 96 *FRFs* given by the ratios of the complex amplitudes of the sound pressures, $p_i(\omega_k)$ and the complex amplitude of the radial force exerted on the box, $f_r(\omega_k)$. The measurements were performed at points located at the centre of a $8 \times 8 = 64$ regular mesh of quadrilateral elements on the vertical surface and at the centre of a $4 \times 8 = 32$ mesh of quadrilateral/triangular elements on the horizontal top surface as described in Section 7.1. Thus, recalling Equation (6.22) the time-averaged total sound power radiation has been derived assuming plane waves and using Equation (4.46) and Equation (4.50). As done for the baffled cylinder, the integration over the radiating surface Ω_a has been substituted into a Riemann mid-point integration over the mesh of elements N_a the surface has been divided into obtaining a form similar to Equation (6.23) but taking into account the different dimensions of each area as represented in Figures 60b.

Moving to the results, the sound power derived using Equation (7.11) from camera, laser and microphones measurements (the latter one taken accordingly to the grid of points described in Section 7.1) are now presented. As anticipated above, the derivation of the sound power from camera and laser data was carried out using the nodal points of the mesh on the surface of the box S_b, as depicted in Figure 55. The orange line in Figure 62, shows the narrow-band spectrum based on the camera measurements whereas the blue line in Figure 63 shows the narrow-band spectrum based on the laser vibrometer measurements. The violet line in both Figures show the narrow-band spectrum derived from the measurements in the semi-anechoic room taken with the array of microphones using Equation (6.23) readapted to account for the dimensions of each measurement area. In general, the spectra obtained from the vibration measurements on the box agree well with those derived from the acoustic measurements.

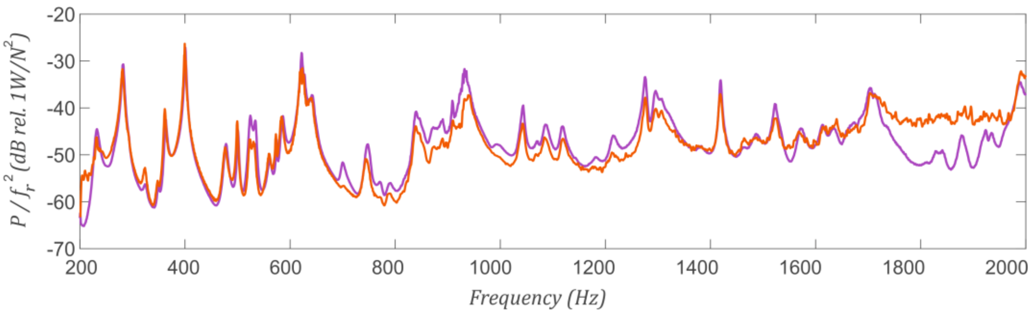


Figure 62: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and calculated from measurements with camera (orange line).

The spectrum obtained from the *camera-based* measurements (orange line in Figure 62) shows only minor deviations compared to the spectrum derived from the acoustic measurements in the anechoic room (violet line). In particular, the agreement is very good over most of the frequency range, with discrepancies appearing mainly around $1000\ Hz$ and at frequencies close to the upper limit of the range considered (as already observed in Chapter 5 for the kinetic energy spectrum). These differences, however, remain relatively small and correspond to sound power radiation levels that, as will be discussed below, exert only a limited influence on the third-octave representation. The spectrum obtained from the laser vibrometer measurements (blue line in Figure 63) also shows very good agreement with the reference acoustic spectrum (violet line).

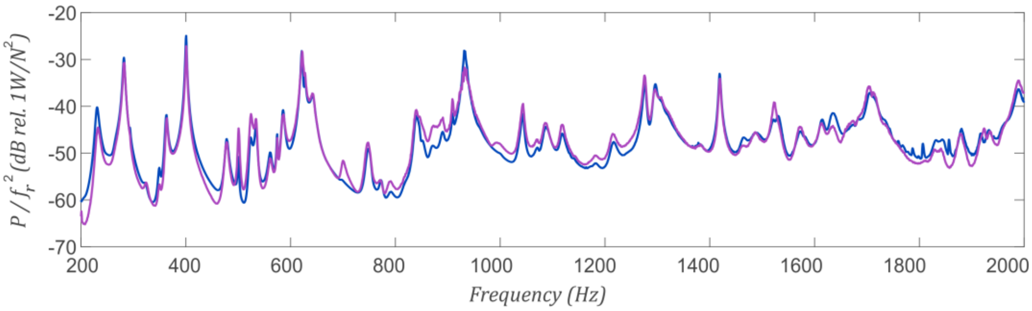


Figure 63: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and calculated from measurements with laser vibrometer (blue line).

As expected, the *laser-based* data present slightly sharper resonance peaks, reflecting the higher spatial resolution of the vibration measurements. A small mismatch can be observed in the frequency band around $900\ Hz$, but overall, the deviations remain limited. When comparing the spectra in third-octave bands of the sound power levels, calculated using Equation (4.53), as illustrated in Figure 64 (orange bars from the camera measurements, blue bars from the laser measurements, violet bars from the acoustic measurements), the consistency among the three methods becomes even clearer. With the exception of the $630\ Hz$ band, which shows very close agreement, the differences remain within $1-2\ dB$, a range generally acceptable for engineering applications. Nevertheless, it should be noted that the acoustic-based measurement relies on a relatively sparse grid of microphone positions and on the quality of the semi-anechoic room, which may be affected by limited absorption in specific frequency bands and by flanking noise contributions. In contrast, the vibration-based approaches rely on dense spatial sampling of the surface velocity field (all mesh nodes reported in Table 3), which are inherently less sensitive to background acoustic noise, unless such noise is strong enough to induce measurable structural vibrations.

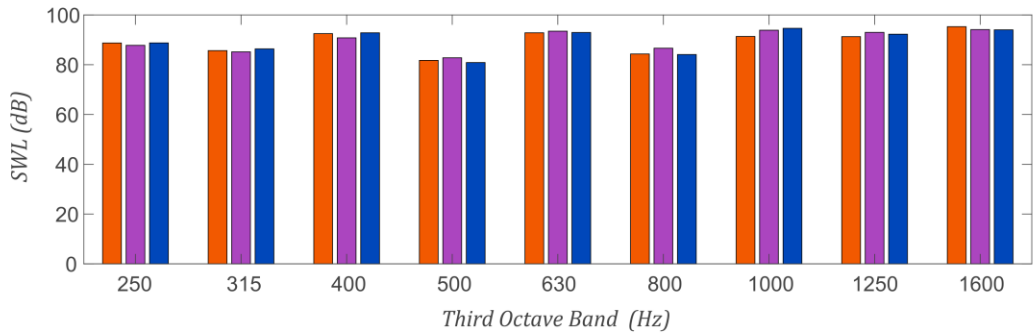


Figure 64: third-octave band spectrum of the total sound power radiation from measurements with microphones (violet bars) and calculated from measurements with laser vibrometer (blue bars) and camera (orange bars) – $dB\ ref.\ 10^{-12}\ W$.

The *camera-based* measurements are, in principle, affected by discretization errors related to pixelization, which depend on both the temporal and spatial resolution of the device. In this study, a high-speed camera with sufficiently high spatial resolution was employed, allowing an accurate reconstruction of the vibration field. Importantly, pixelization-induced errors are expected again to exhibit a random character. Therefore, when the reconstructed vibration field is integrated through the Kirchhoff–Helmholtz formulation to compute the radiated sound power, these errors tend to cancel out, resulting in a more reliable estimate of the overall sound radiation.

It is also worth noting that third-octave band spectra are commonly presented using *A-weighting*, in order to better reflect the frequency-dependent sensitivity of human hearing. For this reason, Figure 65 reports the same third-octave spectra shown in Figure 64, but expressed in terms of *A-weighted levels*. The application of the *A-weighting* curve tends to emphasize the contributions in the mid-to-high frequency range, while attenuating the low-frequency content, which is less relevant from a perceptual standpoint. In the present case, the *A-weighted* spectra preserve the overall agreement between the different measurement approaches, with differences between acoustic- and vibration-based methods remaining within a narrow margin. This representation is particularly useful for practical noise assessment, since it provides a more direct link between the measured or predicted sound power and the expected subjective perception of the emitted noise.

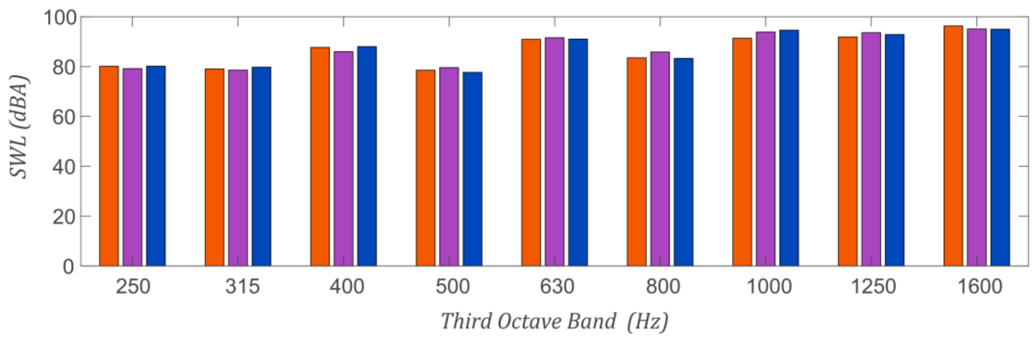


Figure 65: third-octave band spectrum of the total sound power radiation from measurements with microphones (black bars) and calculated from measurements with laser vibrometer (blue bars) and camera (red bars) – $dB\ ref. 10^{-12}\ W$ in dBA.

For completeness, beside the proposed approach based on the integration of the sound intensity over the radiating surface S_{b} (see Figure 60a), the total sound power radiation has also been derived with the classical approach (following *ISO 3744:2010* standards [54,105]), which depend on the integration of the sound intensity over a large surface Ω_a that encloses the radiating object, that is a large cylinder in this case (see Figure 60b). Thus, recalling Equation (4.45) and considering the fact that, in the far-field, the sound waves locally resemble a plane wave over the measurement surface Ω_a such that the particle velocity has radial direction given by Equation (4.46), the instantaneous total sound power radiation considering time-harmonic sound waves results expressed by Equation (4.50).

In principle, the measurement surface should enclose the flexible box, thus it should be a cylinder with large diameter D_S. In this study, a cylinder with finite length H_S has been considered. As done above for the integration over the sound radiation surface S_b in Equation (7.11), the total sound power radiation can be expressed in terms of the following integration:

$$\bar{P}(\omega) = \frac{1}{2} \int_{S_b} \mathcal{R}e\{p_m^*(\boldsymbol{x}, \omega)\dot{u}_r(\boldsymbol{x}, \omega)\} dS_b, \tag{7.13}$$

where $p_m(\boldsymbol{x}, \omega)$ is derived from the *BEM* calculation using Equation (7.4).

Assuming plane waves, Equation (7.13) can be discretized in the following form using Riemann midpoint sums:

$$\bar{P}(\omega) = \frac{1}{2\rho c} \sum_{i=1}^{N_a} S_{bi} |p_m(\omega)|^2. \tag{7.14}$$

Here, S_{bi} represents the $i - th$ area of the measurement surface Ω_a which is divided into $N_a = 96$ elements.

Moving again to the results, the sound power derived using Equation (7.14) from camera, laser, and microphone measurements are now presented. In this case, the derivation of the sound power from camera, laser, and microphone data was performed using the mesh of $N_a = 96$ points depicted in Figure 60b, following the classical formulation in which the sound radiation is integrated over a far-field measurement surface Ω_a. The orange line in Figure 66 shows the narrow-band spectrum obtained from the camera measurements, whereas the blue line in Figure 67 corresponds to the narrow-band spectrum based on the laser vibrometer measurements. In both cases, the violet lines represent the narrow-band spectra derived from the semi-anechoic room measurements with the microphone array, using Equation (7.14). Also in this case, the spectra obtained from the vibration measurements show good agreement with those derived from the acoustic measurements, although some discrepancies remain visible around $750 - 900\ Hz$ for the laser data and across a broader frequency range near $1000\ Hz$ for the camera data. Overall, the spectra presented in Figures 66 and 67 show a somewhat reduced agreement with the reference acoustic spectrum compared to those in Figures 62 and 63. This suggests that the derivation of the total sound power radiation from the integral of the sound intensity over the surface of the box S_b, using Equation (7.11), provides a more accurate estimate than the classical approach based on the integration of sound radiation over a far-field measurement surface Ω_a. This difference can be mainly attributed to the fact that, in the former case, the integration is approximated by a Riemann sum over a smaller surface ($S_b \ll \Omega_a$) with a significantly larger number of points ($N \gg N_a$). In fact, the derivation of the total sound power radiation from the integral of the sound intensity over the surface of the box S_b is performed on a grid of $N = 1206$ points discretising a surface with the same dimensions as the structure. By contrast, the classical approach is based on the integration over a surrounding surface Ω_a, discretised using $N_a = 96$ points, which encloses the entire structure and is located sufficiently far from it to avoid near-field effects.

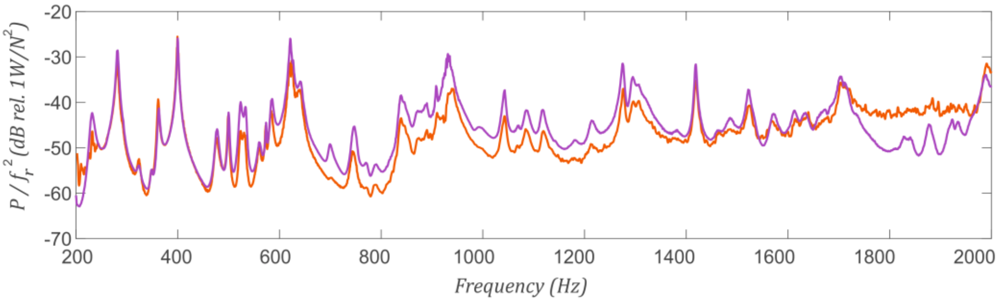


Figure 66: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and reconstructed at microphones positions from measurements with camera (orange line).

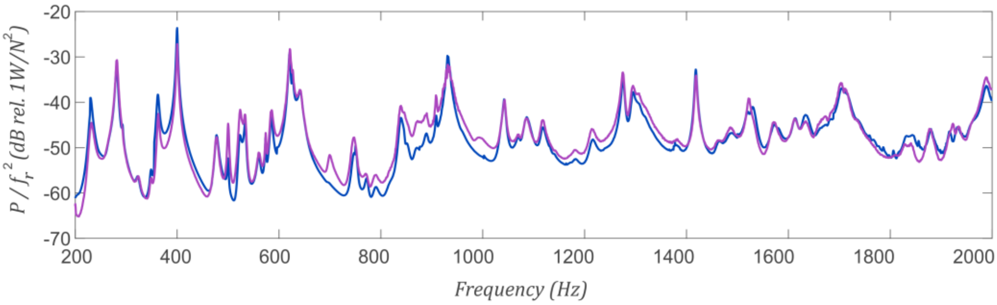


Figure 67: narrow band spectrum of the total sound power radiation per unit force excitation from measurements with microphones (violet line) and reconstructed at microphones positions from measurements with laser vibrometer (blue line).

For the sake of completeness, the one-third-octave band spectra of the sound power levels are also reported in Figure 68, where the orange bars represent the sound power levels from the camera measurements, the blue bars those from the laser measurements, and the violet bars those from the microphone measurements.

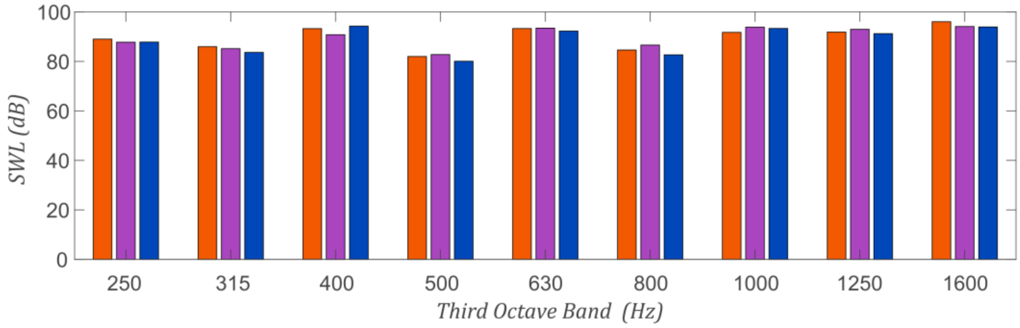


Figure 68: third-octave band spectrum of the total sound power radiation from measurements with microphones (black bars) and reconstructed at microphones positions from measurements with laser vibrometer (blue bars) and camera (red bars) – $dB\ rel.to\ 10^{-12}\ W$.

SUMMARY OF THE CHAPTER

This chapter has addressed the acoustic radiation of a small prismatic box representative of a machine-like structure, characterized by geometric discontinuities and non-smooth boundaries. The analysis focused on the experimental and numerical procedures adopted to estimate the radiated sound field and the total sound power, building upon the vibration data previously introduced in earlier chapters.

The chapter first described the acoustic test rig and measurement setup, detailing the microphone configurations and the spatial discretization employed for the reconstruction of the sound pressure field. The acoustic radiation problem was then formulated using the Kirchhoff–Helmholtz integral equation and solved numerically through a *Boundary Element Method*. Particular attention was devoted to the treatment of geometric discontinuities: kinematic constraints were imposed by enforcing zero normal particle velocity along the box edges. The physical consistency of these constraints was assessed by comparison with the reconstructed normal velocity components, confirming the validity of the adopted boundary conditions. Subsequently, the radiated sound power was evaluated using two complementary approaches. First, the surface-based formulation previously introduced for the cylindrical structure was applied, in which the Kirchhoff–Helmholtz integral is collapsed onto the radiating surfaces and approximated by a Riemann summation over a dense grid of elemental radiators. This approach allows the sound power to be derived directly from the surface vibration data. For further validation, the results were compared with those obtained using the classical ISO-based method, where sound power is estimated through the integration of sound intensity over a surrounding measurement surface. A good agreement between the two methods was observed, supporting the applicability of the proposed formulation to machine-like structures with complex geometries.

Overall, the chapter demonstrates that the adopted acoustic measurement and modelling framework is suitable for estimating radiated sound power in the presence of structural discontinuities, in both narrow-band and one-third-octave band formulations. These results further support the extension of the proposed methodology from idealized laboratory structures to more realistic engineering components.

8
VIBRATION MEASUREMENTS USING EVENT CAMERAS

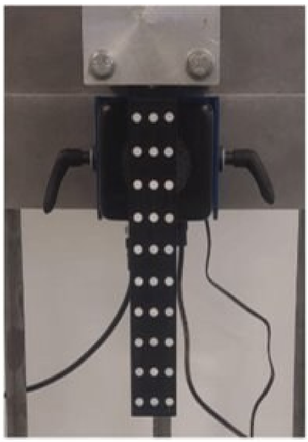
This chapter examines the use of synchronized event cameras to measure the vibration field of a cantilever beam, focusing on its first two flexural resonance modes. Compared to the approaches adopted in the previous chapters, which relied on a single high-speed frame-based camera, the use of event-based sensors represents a methodological advancement aimed at extending the measurable frequency range while significantly reducing the amount of acquired data. Building on the theoretical framework introduced in Chapter 3, the methodology integrates *event-based* vision, intensity frame reconstruction, marker tracking, and multi-view triangulation. After describing the test rig and the calibration of the cameras, intensity images are reconstructed from the event streams through the *E2VID* neural network, enabling reliable identification and tracking of surface markers. The reconstructed three-dimensional trajectories are validated against laser Doppler vibrometer measurements using two error metrics: the absolute error at each marker location and the root mean square deviation over the entire grid. The chapter first addresses the analysis of the first flexural mode [68,69], where both *two-* and *four-camera* configurations are examined in order to assess the influence of the acquisition setup. The discussion then extends to the second flexural mode, which is investigated using the complete *four-camera* configuration.

8.1 TEST STRUCTURE: CANTILEVER BEAM

The test structure employed for the analysis of the first and second flexural mode shapes is a cantilever beam. These deflection shapes correspond to the two lowest resonance frequencies and are detected using a maximum of four synchronized event cameras. The beam, manufactured from steel, has its main physical properties reported in Table 5. As illustrated in Figure 69a, the surface of the beam is discretized into a uniform grid of 10 × 3 points, each identified by a circular marker. The spatial coordinates of these markers are expressed with respect to a reference system (O, x, y, z), positioned near the top right corner of the beam and offset by $5\ mm$ from its longest edge.

Table 5: cantilever beam physical properties

Parameter	**Cantilever Beam**
Length	$L = 210\ mm$
Width	$h = 30\ mm$
Thickness	$s = 2\ mm$
Density	$\rho = 7850\ kg/m^3$
Young's modulus	$E = 21 \times 10^{10}\ N/m^2$
Poisson ratio	$\nu = 0{,}3$
Position of the shaker force excitation from the clamp	$y = 45\ mm$
Grid of measured points first mode	9×3
Grid of measured points first mode	8×3



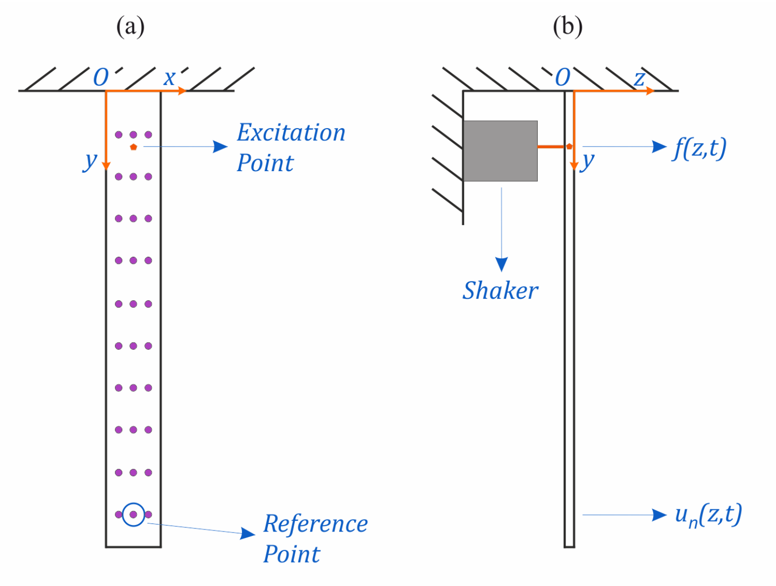


Figure 69: sketch of the experimental setup a) front view of the cantilever beam with reference point, b) lateral view with excitation position and main direction of the displacements

The excitation is provided in bending by an electrodynamic shaker, connected to the beam through a *stinger*, as shown in Figure 69b. In addition, Figure 69a highlights the specific marker chosen as the reference point for representing the time-harmonic response discussed in Section 8.7. The cantilever beam was selected as the test structure owing to its well-defined boundary conditions, geometric simplicity, and the availability of analytical solutions for its resonance frequencies and mode shapes. These characteristics make it a standard benchmark in vibration analysis and an ideal case study for validating the novel proposed measurement approach with event cameras.

8.2 CANTILEVER BEAM FLEXURAL MODES

Considering a uniform beam, composed of linear, homogeneous, isotropic and elastic material, in the free vibration case, the transverse equation of motion can be written as [58,91,92,99]:

$$EI\left(\frac{\partial^4 u_n}{\partial y^4}\right) + \rho A \frac{\partial^2 u_n}{\partial t^2} = 0. \tag{8.1}$$

Here $A = hs$ represents the area of the cross section. Moreover, Equation (8.1) can be rewritten considering complex notation in:

$$\frac{\partial^4 U_n}{\partial y^4} - \frac{\omega^2 \rho A}{EI} U_n = \frac{\partial^4 U_n}{\partial y^4} - \lambda^4 U_n = 0, \tag{8.2}$$

where $u_n(x, y, t) = U_n(x, y) e^{i\omega t}$.

For a cantilever beam the following constraints are set to the boundaries [91,92]:

$$\begin{cases} u_n(x,0,t) = 0 \\ \left[\dfrac{\partial u_n}{\partial y}\right]_{(x,0,t)} = 0 \\ M_{yy}(x,L,t) = \left[-EI\dfrac{\partial^2 u_n}{\partial y^2}\right]_{(x,L,t)} = 0 \\ Q_y(x,L,t) = \left[-EI\dfrac{\partial^3 u_n}{\partial y^3}\right]_{(x,L,t)} = 0 \end{cases} \tag{8.3}$$

and after some mathematical manipulations, using Equation (8.2) and the boundary condition in (8.3), the natural modes can be obtained as:

$$U_{nm}(y) = \frac{1}{\lambda_m^2}\left[\frac{\partial^2 U_{nm}}{\partial y^2}\right]_{y=0}\left[C(\lambda_m y) - \frac{A(\lambda_m L)}{B(\lambda_m L)} D(\lambda_m y)\right]. \tag{8.4}$$

Here the coefficients A, B, C, D are defined as follows:

$$\begin{cases} A(\lambda_m L) = \frac{1}{2}(\cosh(\lambda_m L) + \cos(\lambda_m L)) \\ B(\lambda_m L) = \frac{1}{2}(\sinh(\lambda_m L) + \sin(\lambda_m L)) \\ C(\lambda_m y) = \frac{1}{2}(\cosh(\lambda_m y) - \cos(\lambda_m y)) \\ D(\lambda_m y) = \frac{1}{2}(\sinh(\lambda_m y) - \sin(\lambda_m y)) \end{cases} \tag{8.5}$$

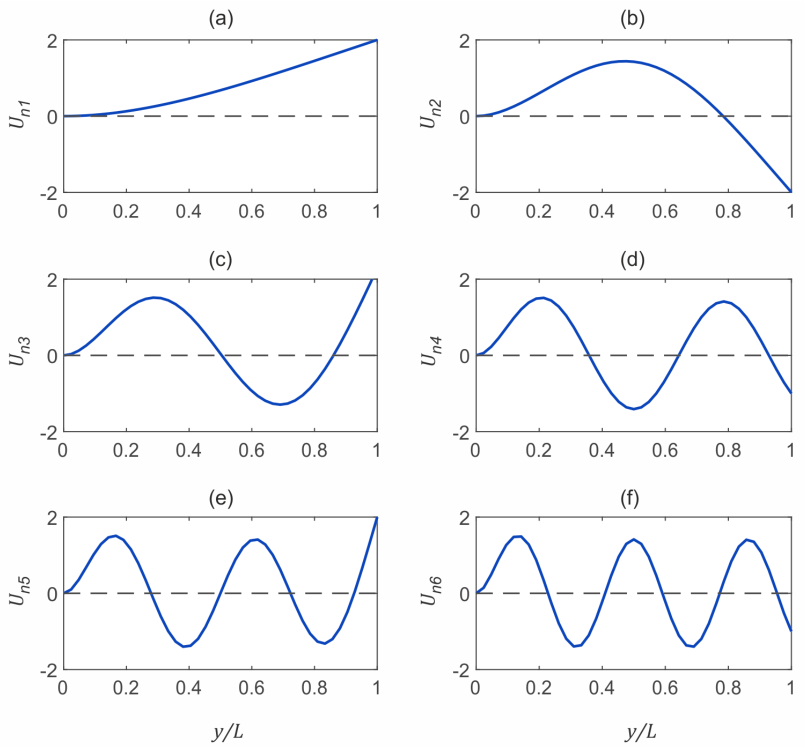


Figure 70: first six deflection shapes of a cantilever beam.

The mode shape is determined by the bracketed quantity in Equation (8.4) and the magnitude of the coefficient $\frac{1}{\lambda_m^2}\left[\frac{\partial^2 U_{nm}}{\partial y^2}\right]_{y=0}$ is arbitrary as far as the mode shape is concerned and is a function of the excitation. Moreover, the natural frequency of each mode can be obtained from:

$$\omega_m = \frac{\lambda_m L}{L^2}\sqrt{\frac{EI}{\rho A}} \tag{8.6}$$

In Figure 70 the first five mode shapes are presented in order to better understand the typical behaviour of the deflection shapes of a cantilever beam. Important to notice are the presence of nodal lines.

8.3 TEST RIGS: EVENT CAMERAS AND LASER VIBROMETER

For the experimental campaign, the setup comprised a maximum of four synchronized event cameras (Figure 71a, b), all oriented toward the central region of the cantilever beam. The cameras were positioned approximately along circular arcs centred on the beam, with a radius of about 50 cm.

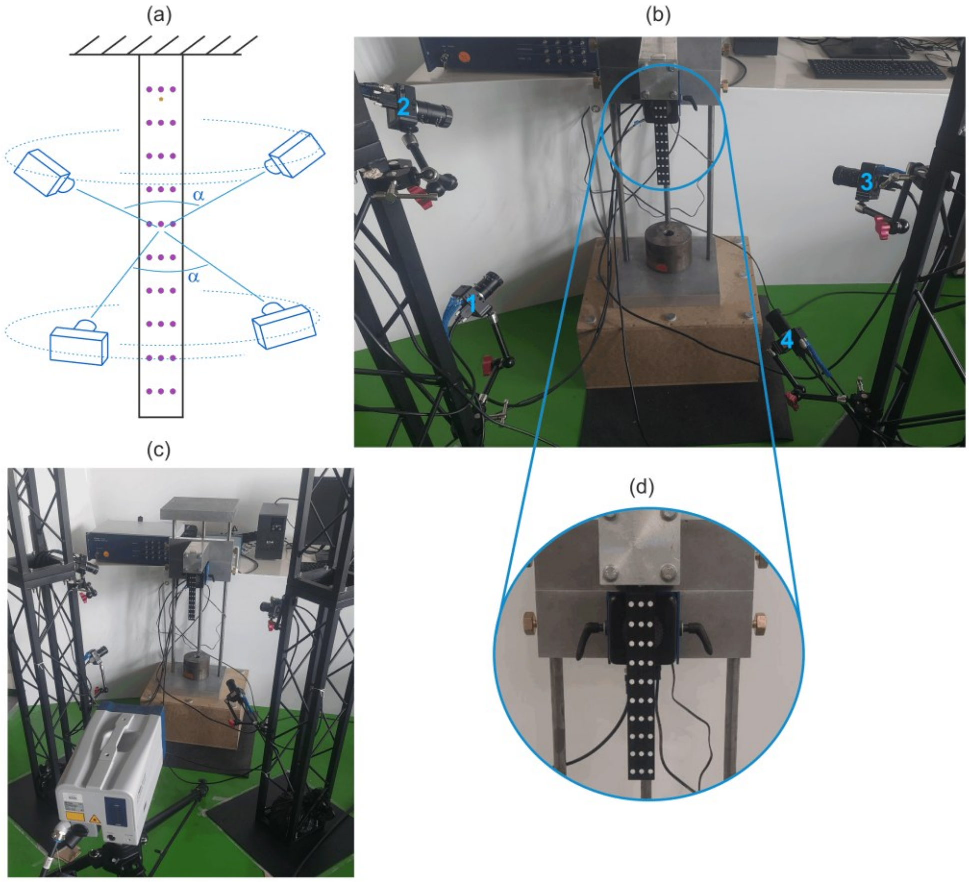


Figure 71: a) sketch of the experimental setup with four event cameras, b) picture of the cantilever beam with 4 event cameras and relative IDs, c) setup including the laser vibrometer d) magnification of the cantilever beam with markers

To enhance spatial coverage and ensure accurate triangulation, the cameras were mounted at two different elevation levels: two cameras at lower heights with a small vertical offset between them, and two at higher positions, again with a slight relative offset. The recordings last for a few seconds after the initial transient has passed, during which tonal measurements are taken. The duration of the recordings depends on the number of periods required for accurate measurement and, therefore, is inversely related to the frequency being analysed. The higher the frequency, the shorter the required recording duration can be. The main specifications of the event cameras employed in this study, which are particularly relevant for dynamic vibration measurements, are summarized in Table 6.

Table 6: main properties of the event cameras employed

Parameter	Event Camera
Model	Inivation DVXplorer
Spatial resolution	$640 \times 480\ pixels$
Temporal resolution	$200\ \mu s$
Typical latency	$< 1\ ms$
Dynamic range	~ 90 dB (3-100k lux with 99.9% of pixels respond to 27.5% contrast) ~ 110 dB (0.3-100k lux with 50% of pixels respond to 80% contrast)
Contrast sensitivity	13% (with 50% of pixels respond) 27.5% (with 99.9% of pixels respond)
Marker diameter	$5\ mm$
Radial distance	$d \cong 500\ mm$
Ground sampling distance (average)	0.56

As illustrated in Figure 71c, the setup also included a scanning laser vibrometer, which was used to acquire reference measurements of the beam transverse vibration field at the predefined grid of marker locations which are shown in Figure 71d. These *laser-based* measurements serve as a benchmark for validating the vibration reconstructions obtained from the event camera system, thereby providing a reliable basis for assessing the accuracy and robustness of the proposed approach.

8.4 CAMERA SYNCHRONISATION VERIFICATION

Before starting the experiments on the cantilever beam and the subsequent data processing, dedicated tests were carried out to verify the synchronization accuracy among the four available cameras. The *DVXplorer* units are hardware-synchronized through daisy-chained connections (Figure 72), with one camera designated as the master. Owing to the built-in synchronization feature, the master camera controls the timestamp reset for all connected units, thereby maintaining temporal alignment with sub-microsecond accuracy.

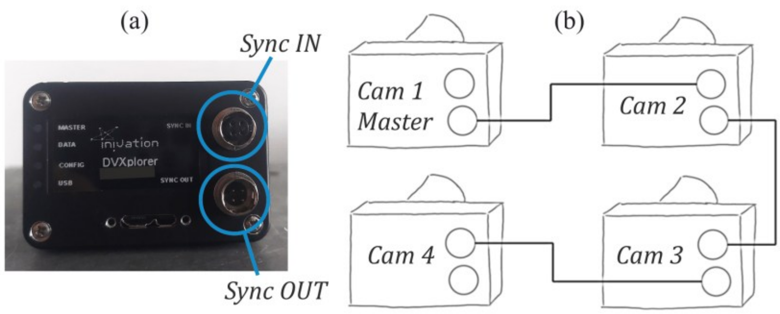


Figure 72: a) main connections of DVXplorer, b) hardware-synchronization through daisy-chained connections.

To validate this synchronization, a dedicated experiment was performed in which all four cameras observed the same portion of a diffusely reflective surface illuminated by a temporally modulated light source. The source was driven by a chirp signal (Figure 73a), designed to generate a known time-varying illumination pattern. This enabled alignment of the event histograms produced by the cameras via a *cross-correlation* procedure. Although the camera timestamps provide microsecond resolution, the histogram-based verification required a coarser temporal binning to ensure statistical reliability. A bin width of $100\ \mu s$ was selected as a compromise, allowing alignment verification with a temporal resolution of $0.1\ ms$.

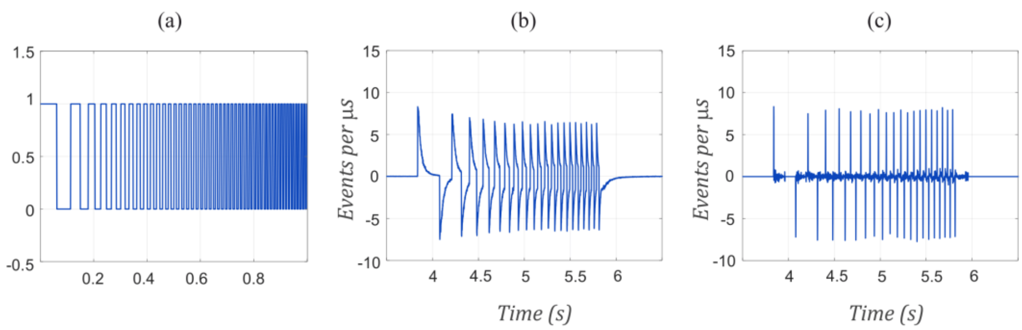


Figure 73: a) chirp signal, b) example of events with sign, c) example of event signal after the AR filtering.

The transformation from raw event streams to histograms followed a systematic procedure. Each event is represented as a triplet (t, p, s), where t is the timestamp, p the pixel location, and $s \in \{+, -\}$ the polarity. To derive a correlation-suitable signal, the time axis was partitioned into intervals of duration T, and the number of events within each interval was counted, yielding the discrete signal $x_E(nT)$. Different strategies were considered for handling event polarity:

- *polarity-agnostic* summation of positive and negative events;
- retention of positive events only;
- preservation of polarity by subtracting negative events from positive events.

Although the *polarity-agnostic* approach already yielded a usable signal, incorporating polarity information sharpened the correlation peaks. The signed-event representation, x_{sign}, was therefore selected, as it provided the most distinct correlation results and was subsequently used in all analyses.

As illustrated in Figure 73b, the resulting event histograms exhibited a non-negligible *"exponential tail"* following the main peak. Ideally, a sharp peak corresponding to the illumination impulse was expected, but a decaying component was observed instead. This effect is attributed to limitations in the event-handling precision of the cameras under high-density stimuli and to internal detection thresholds [66,76], which artificially broaden the histogram and reduce correlation sharpness. Moreover, a residual baseline of spurious events was observed even in intervals without light transitions. To mitigate this, low-amplitude components were thresholded to zero prior to correlation, thereby improving robustness to noise.

To further suppress the exponential decay and enhance alignment precision, a digital filtering strategy was adopted. From a signal-processing perspective, the observed decay can be modelled as an exponential component of the form a^n $(a < 1,\ n > 0)$, corresponding to a pole in the *z-domain*. To cancel this effect, the signal was processed with a first-order autoregressive (*AR*) filter of the form:

$$y(n) = x(n) - \alpha x(n-1), \tag{8.7}$$

where the parameter α (e.g. $\alpha = 0.975$) was identified via model fitting. This filter effectively suppresses exponential decays, thereby improving the temporal resolution of the signal. Figure 73c illustrates the effect of this filtering on the data shown in Figure 73b. After filtering, the peak of the correlation function, refined by parabolic interpolation, was found at a time offset of approximately $0.13\ ms$, indicating the temporal shift required for optimal alignment of the event streams.

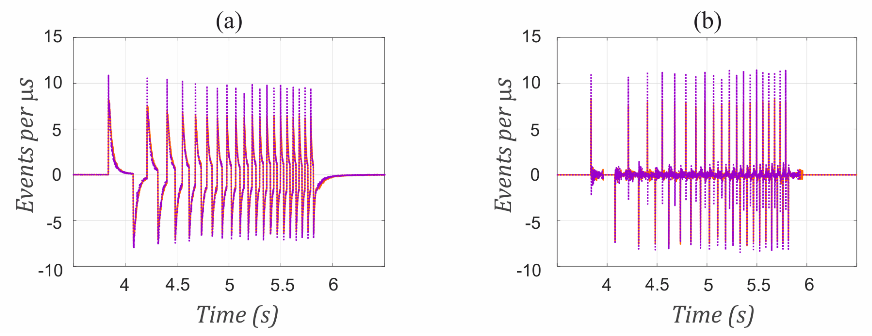


Figure 74: cam 1-2 (a) signals obtained from event sequence with bin width $T = 0.1\ ms$, and (b) signals obtained from (a) after AR filtering.

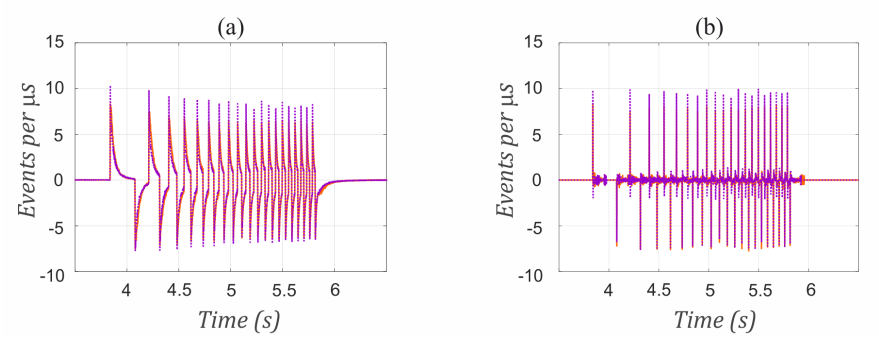


Figure 75: cam 1-3 (a) signals obtained from event sequence with bin width $T = 0.1\ ms$, and (b) signals obtained from (a) after AR filtering.

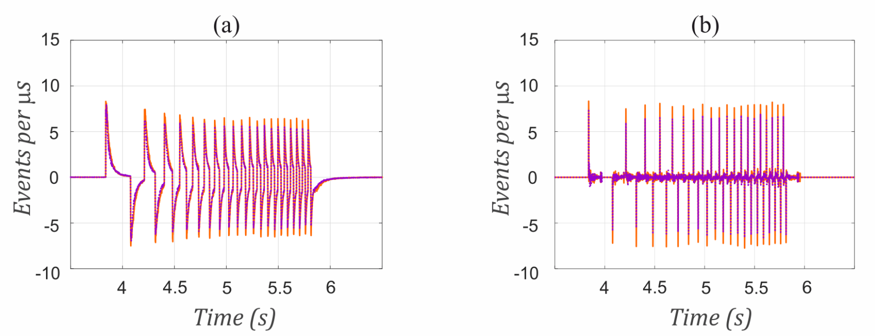


Figure 76: cam 1-4 (a) signals obtained from event sequence with bin width $T = 0.1\ ms$, and (b) signals obtained from (a) after AR filtering.

The verification procedure was repeated for each camera pair involving the master unit (*Master–2*, *Master–3*, *Master–4*). The corresponding results are reported in Figures 74–76, where each panel shows (a) signals obtained from event sequence with bin width

$T = 0.1\,ms$, and (b) signals obtained from (a) after AR filtering. In all cases, the correlation peaks exhibited the same characteristics as the representative example in Figure 73, confirming the robustness of the alignment procedure across all camera pairs. Minor variations in correlation amplitude are attributed to differences in event density between sensors, but the estimated time offsets consistently remained below $0.2\,ms$.

This experiment confirmed that the synchronization accuracy of the camera system lies within a few tenths of a millisecond, a level of precision sufficient to enable high-frequency acquisitions and subsequent triangulation for $3D$ scene reconstruction.

8.5 FRAME RECONSTRUCTION BY *E2VID* NEURAL NETWORK

Although event cameras theoretically encode the complete spatiotemporal information of a dynamic visual scene, the asynchronous data format is fundamentally incompatible with conventional *frame-based* vision post processing. To bridge this gap, two main strategies have been proposed in the literature: direct *event-based* processing [111–114] and the reconstruction of intermediate intensity representations [115–117]. The former, although conceptually elegant, presents significant implementation difficulties and remains relatively underexplored. The latter, which is adopted in this study, reconstructs intensity images directly from the raw event stream, thereby enabling the application of established computer vision techniques. An example of the reconstructed output is shown in Figure 77, where the presence of clearly visible markers demonstrates the feasibility of this approach.

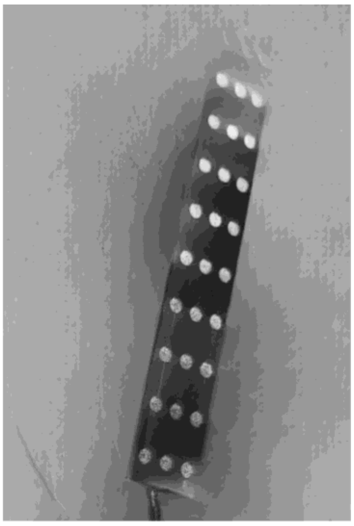


Figure 77: example of reconstructed frame from stream of events.

Within this framework, the reconstruction of intensity images is carried out using *E2VID*, a recurrent fully convolutional neural network specifically designed for *event-based* vision. Its architecture is inspired by the *UNet model* [118] and is organized in an *encoder–decoder* structure with multiple skip connections. *Depth-wise* convolutions enhance computational efficiency, *ConvLSTM units* [119] capture temporal dependencies across event sequences, and bilinear upsampling layers refine the image resolution. Batch normalization and *ReLU* activations are applied throughout the network to ensure stable training and inference. *E2VID* operates in two distinct modes. The first is a *fixed frame-rate mode*, in which reconstructed intensity frames are generated at regular intervals. This mode is particularly suited to *multi-camera* setups and $3D$ triangulation tasks, since it guarantees temporal synchronization across sensors. The second is a *variable frame-rate mode*, in which frames are generated adaptively depending on event density, thereby optimizing reconstruction in scenarios with intermittent motion. The choice of frame rate in both modes must balance event density against image quality: low frame rates can lead to excessive event accumulation and motion blur, whereas high frame rates may result in insufficient event counts, causing loss of structural detail. The reconstruction of intensity images from sparse spatiotemporal event data constitutes an *ill-posed* problem, since events provide only local changes in brightness rather than complete image frames. The transformation of raw event streams into coherent visual representations requires aggregation over short temporal windows. The selection of the window size is therefore a critical parameter, as it directly determines both the temporal resolution and the fidelity of the reconstructed motion. A *well-calibrated* window must satisfy the Nyquist–Shannon sampling criterion to avoid temporal aliasing, while simultaneously ensuring that a sufficient number of events are accumulated to generate intensity images of adequate quality. These reconstructed frames must contain sufficient structural detail to support downstream tasks such as marker tracking and $3D$ triangulation, which are traditionally developed for *frame-based* imagery. In this research thesis, *E2VID* is employed in *fixed frame-rate mode* with overlapping temporal windows, in order to combine temporal synchronization with adequate event density. The window length guarantees the accumulation of a sufficient number of events, while the window shift determines the effective frame rate. As anticipated in Section 8.1, the first two flexural modes are be considered and different overlapping strategies were applied depending on the frequency of the flexural mode under investigation.

CALIBRATION

In this doctoral research, a maximum of four event cameras were employed for vibration measurement and $3D$ reconstruction. As already discussed in Chapter 2, accurate camera calibration is a prerequisite for reliable multi-view analysis. Prior to image acquisition, all four cameras underwent a dedicated calibration procedure. As introduced in Chapter 2, the pinhole camera model [71] is characterized by two distinct sets of parameters: *intrinsic* and *extrinsic*. The *intrinsic* parameters describe the internal geometry and optical properties of the sensor, including focal length, principal point, and lens distortion coefficients. The *extrinsic* parameters, on the other hand, define the spatial pose of the camera within the world coordinate system (encompassing both position and orientation) and are collectively referred to as the exterior orientation. In the present study, the intrinsic calibration of each camera, including radial distortion correction, was performed individually during a preliminary setup phase using the *Sturm–Maybank–Zhang method* [77,78]. For this purpose, multiple views of a planar checkerboard target were acquired independently by each device. These images served as input to the calibration algorithm, from which the intrinsic parameters were obtained. The results are summarised in Table 7.

Table 7: calibration intrinsic parameters of all four event cameras

Camera 1		**Camera 2**	
Focal Length *u* direction	$\alpha_{u,1} = -1241$	Focal Length *u* direction	$\alpha_{u,2} = -1395$
Focal Length *v* direction	$\alpha_{v,1} = -1244$	Focal Length *v* direction	$\alpha_{v,2} = -1403$
Distortion	$\kappa_{0,1} = -0{,}17$	Distortion	$\kappa_{0,2} = -0{,}31$
Ground Sample Distance (GSD) $[mm/pixel]$	$min = 0{,}642$ $max = 1{,}13$	Ground Sample Distance (GSD) $[mm/pixel]$	$min = 0{,}585$ $max = 1{,}26$
Camera 3		**Camera 4**	
Focal Length *u* direction	$\alpha_{u,3} = -1287$	Focal Length *u* direction	$\alpha_{u,4} = -1146$
Focal Length *v* direction	$\alpha_{v,3} = -1285$	Focal Length *v* direction	$\alpha_{v,4} = -1141$
Distortion	$\kappa_{0,3} = -0{,}2$	Distortion	$\kappa_{0,4} = -0{,}3$
Ground Sample Distance (GSD) $[mm/pixel]$	$min = 0{,}891$ $max = 1{,}79$	Ground Sample Distance (GSD) $[mm/pixel]$	$min = 0{,}685$ $max = 1{,}11$

8.6 MARKER TRACKING AND TRIANGULATION

Building on the theoretical framework introduced in Chapter 3, the extraction of $3D$ marker trajectories from event camera data relies on a combination of subpixel tracking and multi-view triangulation. Under the assumption of time-harmonic and small-amplitude vibrations, the marker motion can be well approximated by a sinusoidal model. Once the $2D$ trajectories have been extracted for each marker in each camera view, $3D$ reconstruction is carried out through a classical triangulation procedure comprising the following steps:

a. Initial estimation of extrinsic parameters**:** an initial estimate of each camera extrinsic parameters is obtained via the *Direct Linear Transform* (*DLT*) *method* [71,120]. This approach exploits the known $2D - 3D$ correspondences from the first frame to estimate the camera projection matrix, assuming a pinhole model;

b. Partial bundle adjustment (*BA*): a partial *BA* is then performed to refine the estimated exterior orientations of the cameras, using the *DLT* initialization and the nominal grid coordinates as constraints;

c. Full bundle adjustment: a full bundle adjustment procedure [71] is subsequently applied to jointly optimize both the $3D$ marker positions and the camera extrinsic parameters, minimizing the reprojection error across all views and time steps.

It is important to emphasize that, due to the intrinsic operating principle of event cameras, where data are generated only in response to changes in scene brightness, capturing a static reference frame is not straightforward. To address this limitation, the reference position of each marker is estimated by averaging its tracked coordinates over time, under the assumption of symmetric harmonic motion about the equilibrium point. This methodology, which integrates *event-based* vision, intensity-frame reconstruction, subpixel tracking, and multi-view triangulation, enables high-fidelity vibration analysis with substantially reduced data throughput. Such an approach is particularly advantageous for dynamic structural monitoring applications, where efficiency and accuracy must be balanced. This part of doctoral research is focused on the first two flexural modes of a cantilever beam, so for this study two separate experimental tests were conducted, each corresponding to one of the first two flexural resonance frequencies. The first deflection shape was identified at approximately $32\ Hz$, while the second occurred at around $190\ Hz$.

As introduced in Section 8.5, following the acquisition of both beam vibration data and chessboard calibration sequences, intensity images were reconstructed using the *E2VID* neural network [121]. The sampling threshold was set to $1000\ Hz$ for the first flexural mode and to $1600\ Hz$ for the second. Although these values exceed the minimum required by the Nyquist criterion, they were deliberately selected to explore the temporal resolution limits achievable with the proposed approach. It is worth noting that beam-like structures typically exhibit low modal overlap, which increases proportionally to the square root of the excitation frequency, $\sqrt{\omega}$, as discussed in [97]. Consequently, at each resonance frequency the time-harmonic response is predominantly governed by the corresponding mode shape. Moreover, under the assumption of small-amplitude displacements, the structural response can be accurately approximated as linear, thereby enabling a reliable modal decomposition of the vibration field. In each rectified frame, candidate marker centres were identified using template matching, implemented via normalized *cross-correlation* between the reconstructed image and a predefined template patch. However, due to the presence of noise and illumination artefacts, several false positives were observed (Figure 78). To address this issue and robustly associate the detected points with their theoretical positions in the predefined grid, the assignment algorithm described in Section 3.1 was employed.

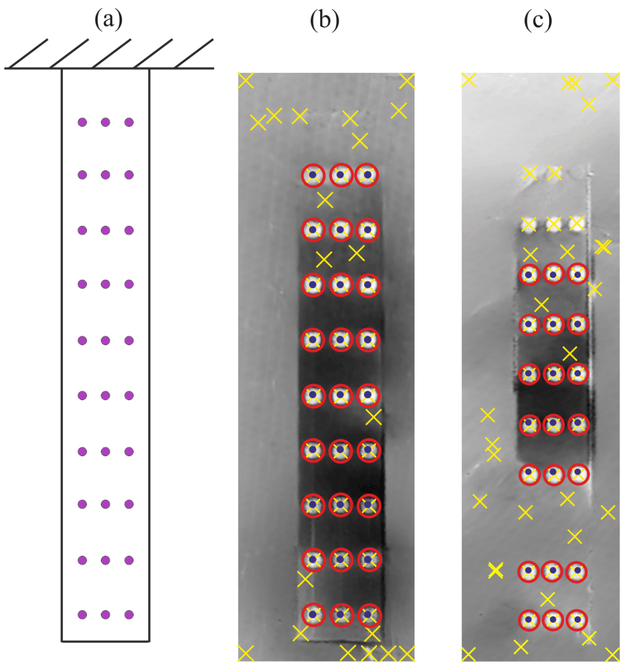


Figure 78: a) sketch of the cantilever beam, b) and c) rectified image in which circles are detected. The violet dots depict the nominal positions (set A), while yellow crosses represent spurious circles that have been detected by template matching but not matched (set B). The red circles are the remaining ones after the assignment problem. b) first flexural mode, c) second flexural mode

8.7 FIRST FLEXURAL DEFLECTION SHAPE

For the first flexural deflection shape, around 32 Hz, where the event density was relatively high, a zero-overlap configuration was sufficient to produce reconstructions with clearly defined markers, enabling robust tracking. The experiment was conducted with all four cameras operating simultaneously; however, in order to perform a parametric study, subsets of the dataset were subsequently extracted to simulate acquisitions with only two cameras. The tracking process yielded $2D$ trajectories for each of the 27 markers (grid of 9×3). Figure 79 shows the tracked paths for each camera (denoted with *ID* numbers in Figure 71b) which are superimposed on a single reconstructed frame. It is evident that the displacements are minimal, on the order of a few pixels.

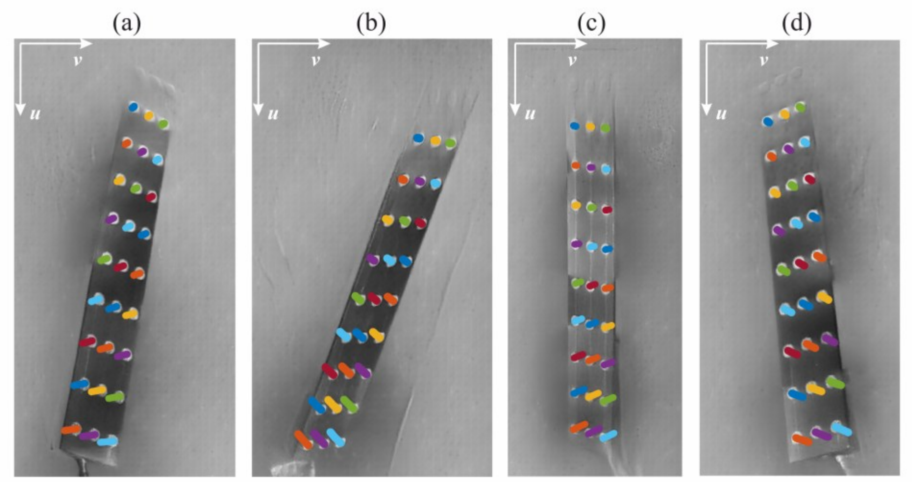


Figure 79: trace of markers for each camera with image coordinate system with v vertical and u horizontal components, a) camera 1, b) camera 2, c) camera 3, d) camera 4

This observation is corroborated by the plots in Figure 80, which display the time-history of the u and v coordinates for the tracked points. The signals derived from the analysis of the events recorded by the four cameras exhibit sinusoidal trends consistent with the expected motion of the beam. The thicker line highlights the trajectory of the reference marker (the one indicated in Figure 69a), which exhibits the largest oscillation amplitude. In the vertical component v, a maximum displacement of approximately 7 pixels is observed, while in the horizontal component u, the peak displacement reaches about 6 pixels. The differences observed among the four event cameras (Figure 80 a–d) reflect their distinct orientations and, consequently, their varying visibility conditions. The signals derived from the analysis of the events recorded by the four cameras exhibit sinusoidal time-histories consistent with the expected motion of the beam.

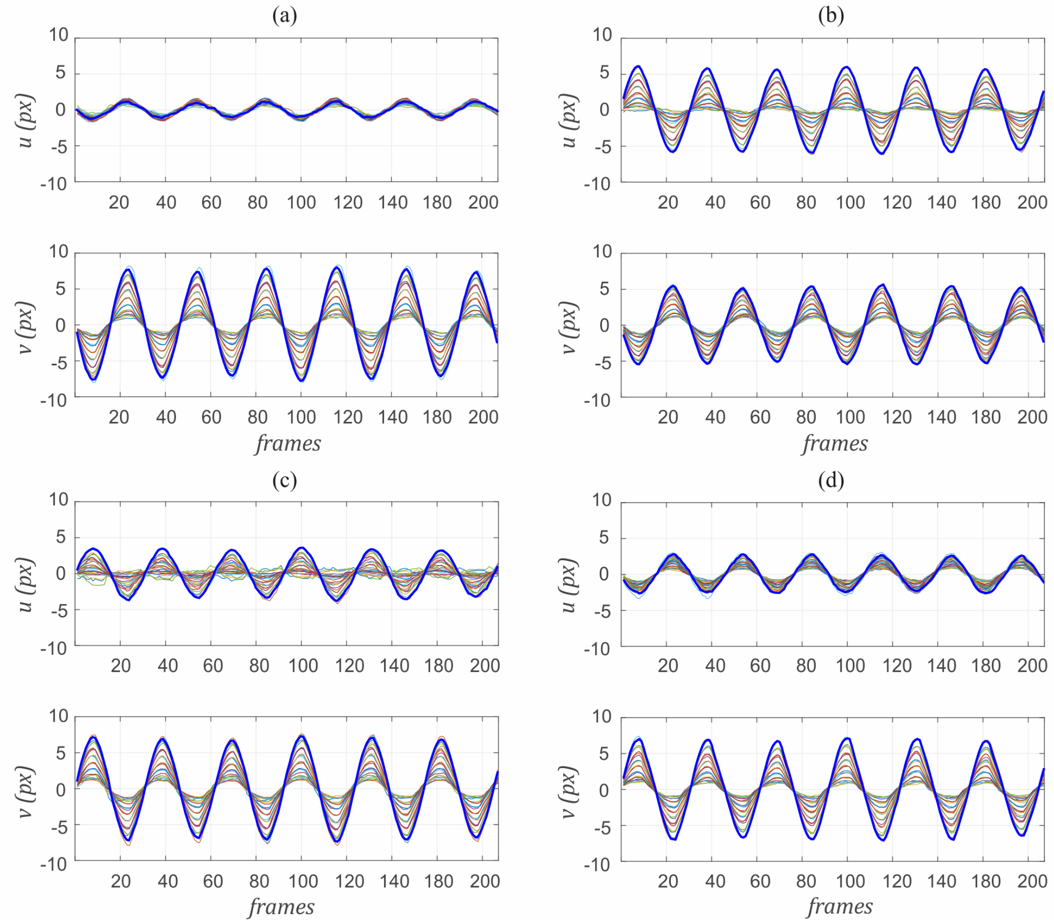


Figure 80: horizontal and vertical displacements of 27 markers in the image coordinate system. Thicker line for the reference point at the tip. a) camera 1, b) camera 2, c) camera 3, d) camera 4

The first approach was based on considering pairs of cameras (*1-2*, *1-3*, *1-4*) allowing for a systematic comparison of the reconstructed $3D$ trajectories across different stereoscopic configurations. Thus, marker trajectories were processed via triangulation, as outlined in Chapter 3, to obtain the corresponding $3D$ vibration fields reported in Figure 81 a-c. Subsequently, the full configuration including all four cameras (*1–2–3–4*) was analysed (Figure 81d) to better understand the improvements obtained adding cameras.

While all configurations provide qualitatively similar motion profiles, particularly along the Z-axis, that is the principal direction of the flexural motion, the full multi-view setup in Figure 81d results in superior reconstruction quality. Specifically, the trajectories obtained using all four cameras exhibit reduced noise and improved spatial consistency, particularly along the X and Y axes, where *single-pair* reconstructions tend to suffer from higher variance.

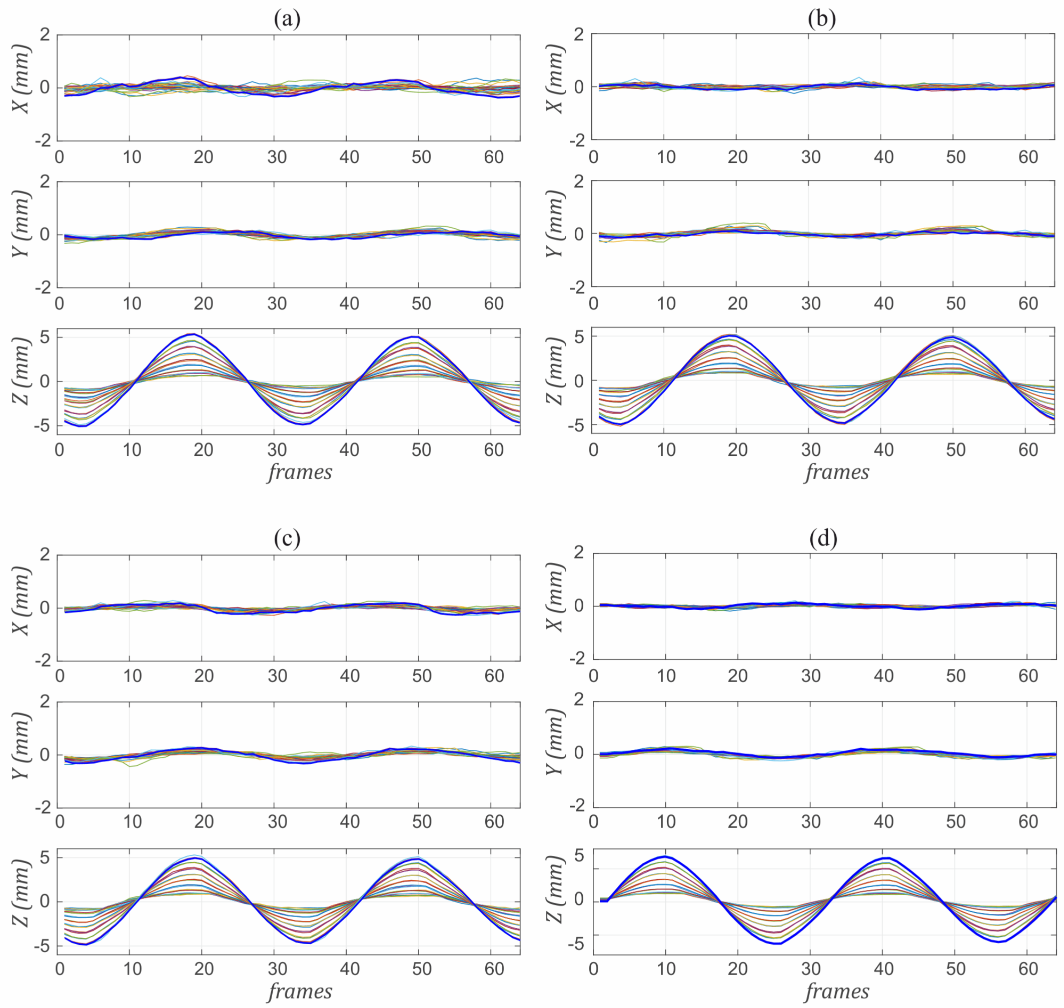


Figure 81: X, Y, Z displacements for the first flexural mode in millimetres for all 27 markers. Thicker line for the reference point at the tip. a) triangulation with cam 1-2, b) triangulation with cam 1-3, c) triangulation with cam 1-4, d) triangulation with cam 1-2-3-4

Notably, the *two-camera* configuration also delivers a relatively clean reconstruction, demonstrating that under favourable geometric conditions, accurate results can be achieved even with minimal camera setups. However, this performance is highly dependent on the viewing angle and baseline of the camera pair, as evidenced by the less stable reconstruction in X and Y directions in Figure 81a. Such sensitivity to camera placement can lead to inconsistent results across different movements or experimental setups. In contrast, incorporating more than two cameras, as in configuration (d), enhances spatial coverage and introduces additional geometric constraints, improving triangulation robustness and reducing the influence of suboptimal camera positioning. Therefore, while acceptable results can be obtained from specific *two-camera* arrangements, a multi-camera setup offers a more reliable solution for high-accuracy and consistent $3D$ motion tracking.

DEVIATION FROM LASER VIBROMETER MEASUREMENTS

As anticipated in Section 8.3, the triangulated results obtained from the event-camera system were validated against reference measurements acquired with a laser Doppler vibrometer (*LDV*). To this end, the experimental setup was equipped with a *Polytec PSV-500-A LDV*, as illustrated in Figure 71c. This instrument served as a benchmark for capturing accurately the transverse vibration field of the cantilever beam. Measurements were performed at the grid of 27 predefined marker positions, ensuring direct comparability with the *event-based* reconstructions. For each point, the transverse displacement in the Z-direction was estimated as the average *peak-to-peak* amplitude observed over two complete vibration cycles.

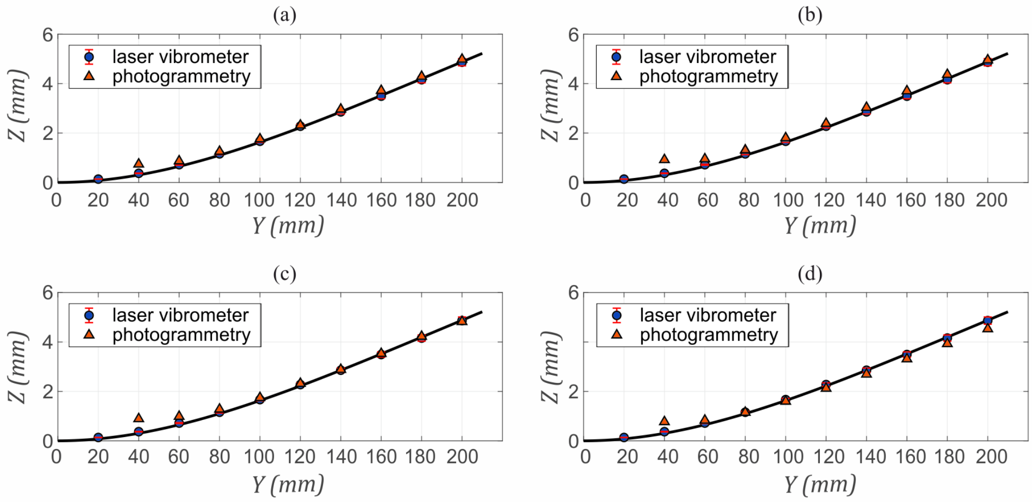


Figure 82: comparison of the average transverse displacements measured with event cameras and laser vibrometer for the first flexural mode. orange triangles for event cameras and blue circles for laser vibrometer. a) cam 1-2, b) cam 1-3, c) cam 1-4, d) cam 1-2-3-4.

This procedure provided a reliable reference dataset against which the *event-based* $3D$ reconstructions could be assessed. The comparative results of the *LDV* reference measurements and the event camera triangulations are reported in Figures 82a–d, corresponding respectively to the following camera configurations: (a) cameras *1–2*, (b) cameras *1–3*, (c) cameras *1–4*, and (d) cameras *1–2–3–4*. The black line in Figures 82a-d depicts the first flexural mode shape of a cantilever beam obtained from the expression:

$$\phi_n(x) = (\cosh(k_n x) - \cos(k_n x)) - \sigma_n(\sinh(k_n x) + \sin(k_n x)), \quad (8.8)$$

where:

$$\sigma_n(x) = \frac{\sinh(k_n L) - \sin(k_n L)}{\cosh(k_n L) + \cos(k_n L)} \tag{8.9}$$

Here k_n is the modal wavenumber, L is the length of the beam and n is the mode. The expression in Equation (8.8) was fitted in such a way as to minimise the mean error with respect to the values measured with the laser vibrometer (blue circles). The agreement between the two approaches is here qualitatively discussed. In all configurations with two cameras, Figure 82a–c, the photogrammetric deformation profiles show good agreement with the laser vibrometer reference. However, the *four-camera* setup (d) yields overall qualitatively better results, as the increased spatial coverage and redundancy of viewpoints reduce the impact of local errors and potential point losses. This ensures a more stable and robust reconstruction, making the *four-camera* configuration the most suitable choice, particularly in more complex scenarios. To better understand the advantages of employing four cameras instead of two, an error analysis was performed, quantifying the reconstruction accuracy as a function of the number of cameras employed. The first error considered is an absolute error defined as:

$$E_{event-laser} = \left| \frac{1}{3} \sum_{i=1}^{3} u_{nc}(x_i, y_i, t_k) - u_{nv}(x_i, y_i, t_k) \right| \tag{8.10}$$

Here, $u_{nc}(x_j, y_j, t_k)$ and $u_{nv}(x_j, y_j, t_k)$ are the transverse displacements derived respectively from the camera and the laser vibrometer measurements at position (x_j, y_j) and time t_k., the i-index represents points on the same row.

In Figure 83, the absolute error calculated using Equation (8.10) is reported as histograms, while the blue shaded area represents the uncertainty range of the laser vibrometer which according to the datasheet is about 3% of the vibration amplitude (this error is also displaced in Figure 82a-c at the marker points with vertical red bars). Consequently, all orange bars falling within the blue region lie inside the measurement uncertainty of the laser vibrometer and can therefore be considered consistent with the benchmark measurement. Moreover, with the error plots in Figure 83(a–d) it is noticeable that the triangulated results in Figure 83(d), obtained using all four cameras, exhibited improved robustness because using more than two cameras ensures a better-quality geometric coverage, as anticipated and described above. Furthermore, it can be observed that configurations with two cameras can still provide satisfactory results; however, in cases (b) and (c), the four measurement points closest to the clamped end of the cantilever beam exhibit larger errors compared to cases (a) and (d). Moreover, it is important to recall that the row of markers positioned closest to the clamped boundary exhibits transverse motion of less than $0.8\,mm$, making displacement detection by the event cameras particularly challenging and highly dependent on camera orientation. The behaviour of the two cameras configuration is particularly interesting, since in case (c) the trend for the points with larger displacements is instead improved. Such an outcome highlights that the viewing angle covered by only two event cameras is not sufficient to guarantee a robust representation of the $3D$ geometry. Therefore, increasing the number of cameras proves advantageous, as it allows balancing robustness and accuracy across all reconstructed points, while reducing the strong dependence on the viewing angle and on the specific orientation of the cameras.

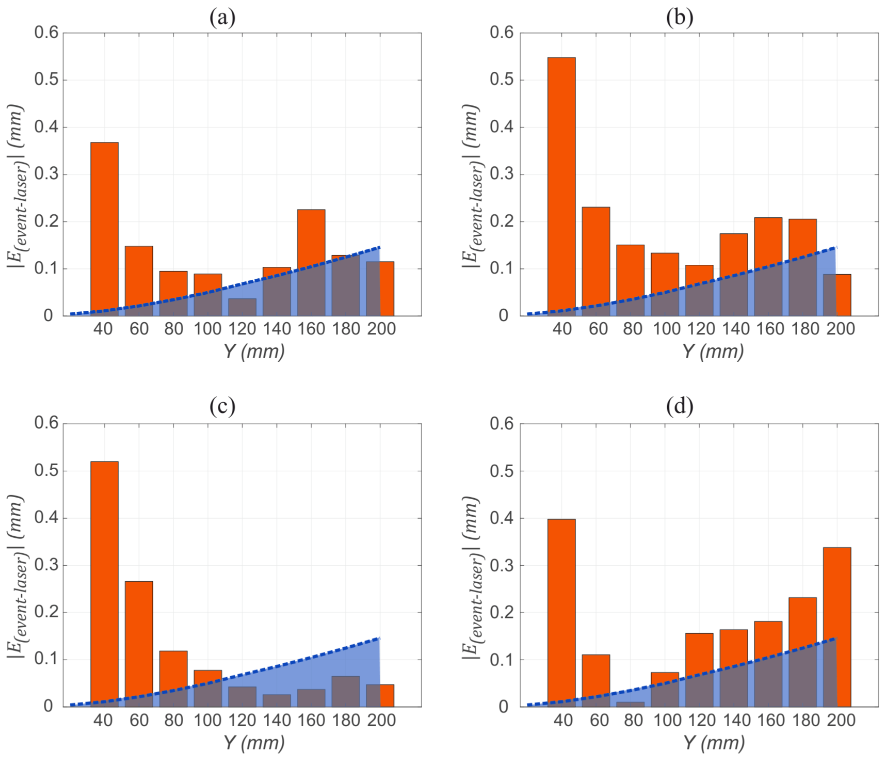


Figure 83: absolute error between camera and laser vibrometer measurements. Blue regions represent the measurement uncertainty of LDV. a) cam 1-2, b) cam 1-3, c) cam 1-4, d) cam 1-2-3-4.

In addition to the absolute error, which focuses on the accuracy at each individual marker position along the Y-axis, an average error was also computed according to the following expression:

$$E_{RMS,u_n} = \sqrt{\frac{1}{N}\sum_{j}\left[w_c(x_j, y_j, t_k) - w_v(x_j, y_j, t_k)\right]^2} \tag{8.11}$$

Where N represents the number of measurement points. The root mean square deviation E_{RMS,u_n} between the event-based measurements and those obtained via *LDV* was calculated and reported in Table 8.

Table 8: root mean square error for different configurations of cameras

	Cam 1-2	**Cam 1-3**	**Cam 1-4**	**Cam 1-2-3-4**
$E_{RMS,w}$	0.18	0.25	0.21	0.22

The results reported in this table further highlight the importance of camera orientation. In fact, when considering pairs of cameras, a root mean square error E_{RMS,u_n} of $0.18\ mm$ is obtained for the configuration composed of cameras *1–2*, whereas for the configuration with cameras *1–3* the E_{RMS,u_n} increases to $0.25\ mm$. Moreover, an E_{RMS,u_n} of $0.22\ mm$ is obtained for the four-camera configuration which, although higher than the root mean square error obtained for the *1–2* configuration, provides a more balanced situation. By mitigating the error sensitivity to camera orientation while introducing redundancy, this configuration significantly enhances the robustness of the measurement system, making the *four-camera* setup the most reliable choice overall.

8.8 SECOND FLEXURAL DEFLECTION SHAPE

Having established that the *four-camera* configuration provides greater robustness and overall reliability, the analysis of the second flexural deflection shape is herein conducted using the complete setup comprising all four event cameras. The second flexural mode of the cantilever beam under test is around $190\ Hz$ and in this case, due to the fact that fewer events were generated within the same time span, a 50% overlapping window was introduced.

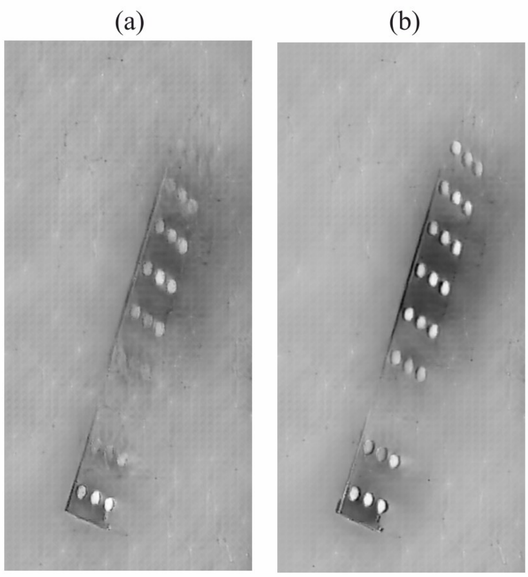


Figure 84: same intensity frame reconstructed using E2VID with a) 0% overlap and b) 50% overlap.

This approach increased the effective event density per frame, thereby improving the visibility of the markers and enhancing tracking reliability. Figure 84 shows the same frame reconstructed with (a) 0% overlap and with (b) 50% overlap, here, the comparison between these two configurations demonstrates how the overlapping strategy can significantly influence the quality of the reconstructed intensity frames and, consequently, the performance of the subsequent vibration analysis.

Again, the tracking process yielded $2D$ trajectories for each of the 21 markers on a grid of 7×3 (one row is coincident with the nodal line of the second deflection shape of the cantilever beam, therefore the displacements are negligible and the row is not detected by the event cameras). Figure 85 shows the tracked paths for each camera which are superimposed on a single reconstructed frame. It is evident that the displacements, as the frequency of the vibration rises, become smaller and challenging to be detected. For the second flexural resonance frequency the tracked paths of each marker are less than 2 pixels. This observation is corroborated by the plots in Figure 86, which display the temporal evolution of the u and v coordinates for the tracked points.

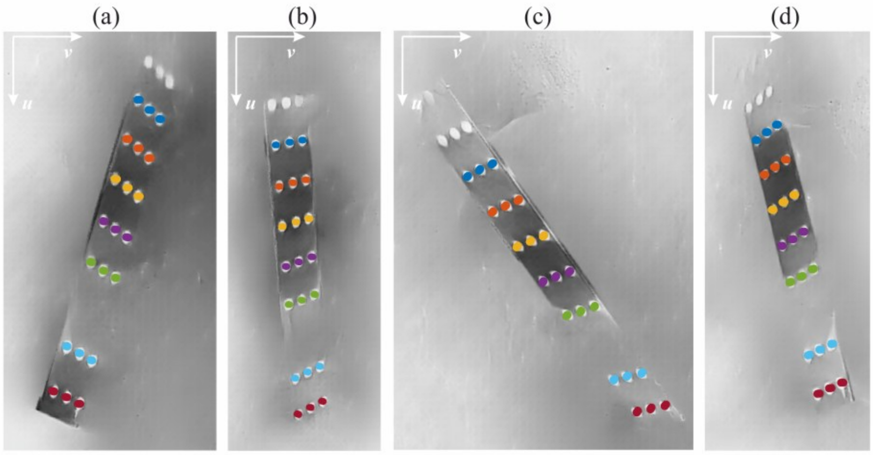


Figure 85: trace of markers for each camera with image coordinate system with v vertical and u horizontal components, a) camera 1, b) camera 2, c) camera 3, d) camera 4

The signals derived from the analysis of the events recorded by the four cameras exhibit sinusoidal trends consistent with the expected motion of the beam also for the second deflection shape here analysed. The thicker line highlights the trajectory of the reference marker (the one indicated in Figure 69a), which should exhibit the largest oscillation amplitude. Figure 86 shows that the vertical component v has a maximum displacement of approximately 1.5 pixels, while in the horizontal component u, the peak displacement reaches about 1 pixels. The differences observed among the four event cameras (a–d) reflect their distinct orientations and, consequently, their varying visibility conditions.

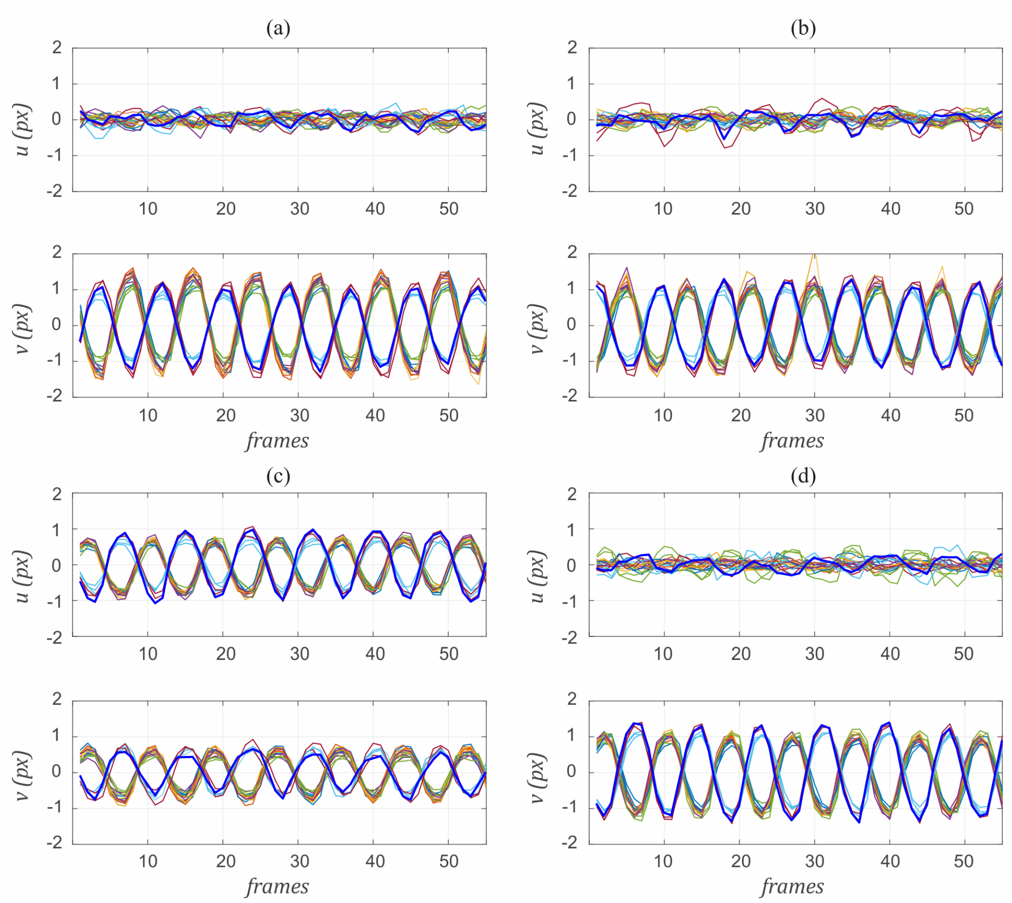


Figure 86: horizontal and vertical displacements of 21 markers in the image coordinate system. Thicker line for the reference point at the tip. a) camera 1, b) camera 2, c) camera 3, d) camera 4

After reconstructing the trajectories of all 21 visible markers in the image sequences from each camera, the triangulation procedure was carried out to obtain their $3D$ coordinates. At higher frequency respect the one previously analysed, the amplitude of the displacements become smaller, and more challenging to be detected and reconstructed. As illustrated in Figure 87, the displacements in the X, Y, and Z directions exhibit sinusoidal time-histories with identical frequency and opposite phase depending on the marker position with respect to the nodal line. In this case too, the frequency of the harmonic response coincides with that of the force excitation exerted by the shaker. The Z component captures the dominant transverse motion associated with the beam flexural response, while the X and Y components represent in-plane displacements. Although the latter are of much smaller amplitude, they still show a coherent, harmonic evolution over time. Figure 87 also highlights with a thicker curve the trajectory of the reference marker located at the tip of the cantilever beam as depicted in Figure 69a.

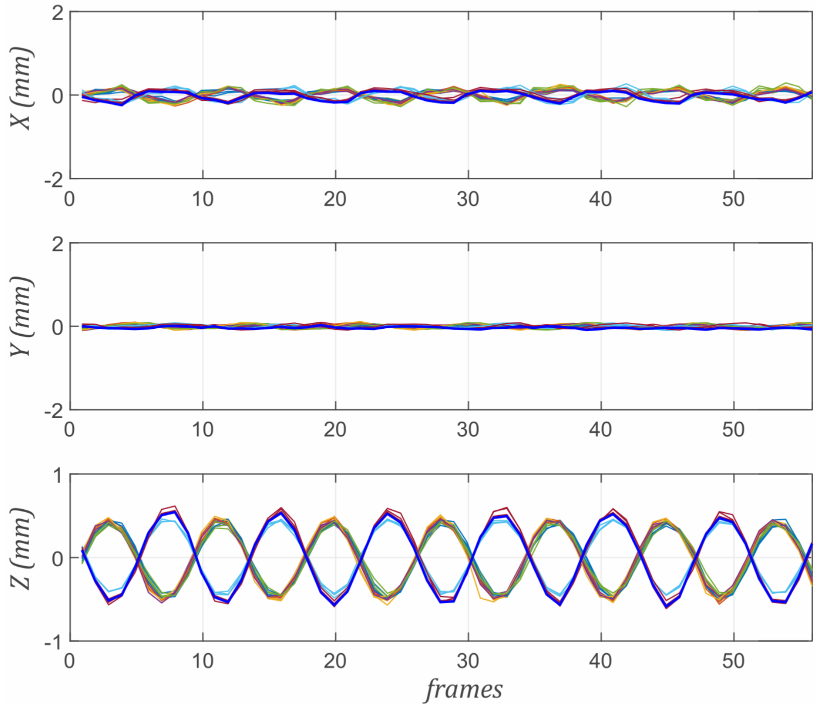


Figure 87: X, Y, Z displacements for the second flexural mode in millimetres for all 21 visible markers. Thicker line for the reference point at the tip

DEVIATION FROM LASER VIBROMETER MEASUREMENTS

As anticipated in Section 8.3 and done in Section 8.7 for the first flexural mode, the triangulated results obtained from the event camera system were validated against reference measurements acquired with a laser Doppler vibrometer (*LDV*). To this end, the experimental setup was equipped with a *Polytec PSV-500-A LDV*, as illustrated in Figure 71d. This instrument served as a benchmark for accurately capturing the transverse vibration field of the cantilever beam. Measurements were performed at the grid of 24 predefined marker positions, ensuring direct comparability with the event-based reconstructions. For each point, the transverse displacement in the Z-direction was again estimated as the average peak-to-peak amplitude observed over two complete vibration cycles. This procedure provided a reliable reference dataset against which the event-based $3D$ reconstructions could be assessed.

The comparative results of the *LDV* reference measurements and the event-camera triangulations are reported in Figures 88. Again, the black line in Figure 88 depicts the second flexural mode shape of a cantilever beam obtained from Equation (8.8) and fitted in such a way to minimise the mean error with respect to the values measured with the laser vibrometer (blue circles).

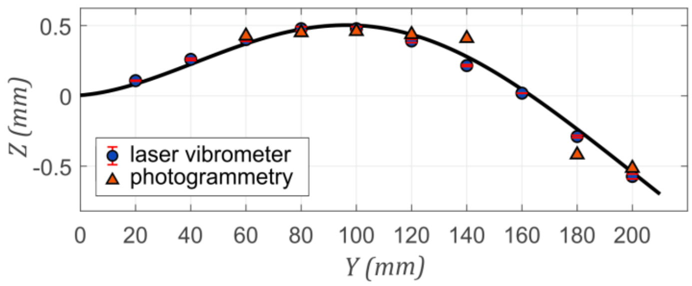


Figure 88: comparison of the average transverse displacements measured with event cameras and laser vibrometer for the second flexural mode. orange triangles for event cameras and blue circles for laser vibrometer.

For the second flexural deflection shape, the row of markers positioned closest to the clamped boundary (in this case the third row of markers from the clamped end) exhibits transverse motion of less than $0.5\ mm$, and the rows before and after the nodal line exhibits transverse motion of less than $0.3\ mm$ making displacement detection by the event cameras particularly challenging. Nevertheless, the use of four distinct viewing angles and the overlapping window set at 50%, enabled a quite accurate reconstruction of the displacements even in these three regions. Also in this case, an error analysis was performed, quantifying the reconstruction accuracy.

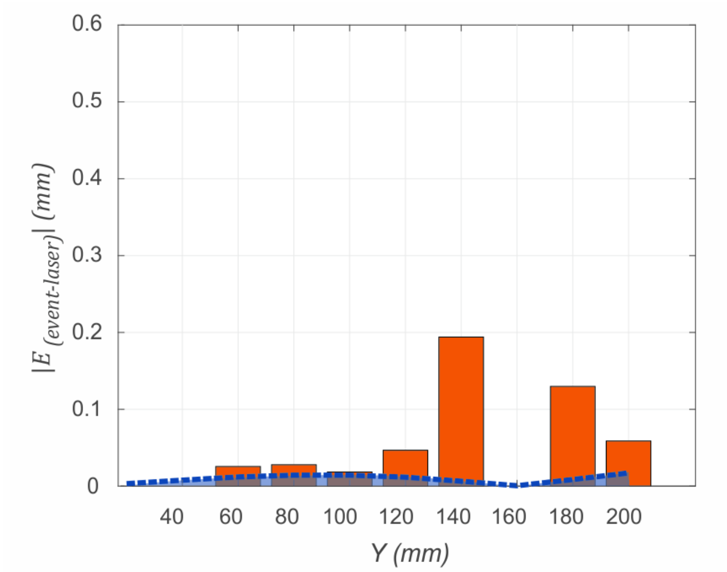


Figure 89: absolute error between camera and laser vibrometer measurements. Blue regions represent the measurement uncertainty of LDV

Again, the first error considered is the absolute error defined in Equation (8.10) and reported in Figure 89 as histograms, while the blue shaded area represents the uncertainty range of the laser vibrometer which according to the datasheet is about 3% of the vibration amplitude (this error is also displaced in Figure 88 at the marker points with vertical red bars). Consequently, all orange bars falling within the blue region lie inside the measurement uncertainty of the laser vibrometer and can therefore be considered consistent with the benchmark measurement. Also in this case, it is notable that the detected markers which exhibit the lowest displacements are affected by the highest absolute error, the critical points are in fact positioned before and after the nodal line. As done for the first flexural mode, in addition to the absolute error, which focuses on the accuracy at each individual marker position along the Y-axis, an average error was also computed according to Equation (8.11). The root mean square deviation E_{RMS,u_n} between the *event-based* measurements and those obtained via *LDV* was calculated and is equal to $0.1\ mm$.

Two additional considerations have to be done: first, the experiments were conducted using event cameras that have relatively low spatial resolution. It is notable that both the precision and quality of the results could be substantially improved through the adoption of higher-resolution event cameras. Second, although the experimental validation focused on the first and second flexural resonance frequency, the methodology is inherently capable of capturing responses at significantly higher frequencies. Given the high temporal resolution of event cameras, the principal limitation at elevated frequencies would be their spatial resolution, as the vibration amplitudes tend to decrease with increasing frequency, making displacement detection more challenging.

SUMMARY OF THE CHAPTER

This chapter focuses on the experimental and methodological development of vibration measurements using an innovative optical approach based on event cameras, distinguishing it from the preceding chapters where a single frame-based high-speed camera was employed. The discussion begins with a description of the test structure, a harmonically excited cantilever beam, and its first two flexural modes. The experimental setup is then detailed, including both the event camera array and a reference laser Doppler vibrometer system, used for validation purposes. A dedicated verification of camera synchronization is presented, confirming sub-millisecond temporal alignment among the four devices. The reconstruction of intensity frames from raw event streams is achieved using the E2VID neural network, enabling robust marker detection and subsequent 3D triangulation. Camera calibration procedures, including the estimation of intrinsic parameters, are also discussed. The chapter further presents the extraction of the 3D deflection shapes for the first and second flexural modes, illustrating the application of the complete photogrammetric pipeline and quantifying the measurement accuracy through comparison with reference LDV data. Parametric analyses highlight the effect of varying the number of cameras from two to four, demonstrating the enhanced robustness and precision achievable with a full four-camera configuration.

Overall, the chapter demonstrates that event cameras, when combined with neural-network-based frame reconstruction, calibration, and multi-view triangulation, provide a reliable and high-temporal-resolution tool for structural vibration measurement, offering a practical alternative to conventional frame-based imaging systems.

9
CONCLUSIONS

This dissertation has addressed two tightly coupled objectives: first, the experimental reconstruction of three-dimensional vibration fields of engineering structures via optical and *event-based* vision, and second the quantitative derivation of the acoustic radiation produced by those vibrating structures, with particular emphasis on the estimation of sound power. The research has therefore bridged structural dynamics, optical sensing, and computational acoustics: *full-field* vibration measurements were used as the primary input for boundary integral and boundary element formulations that predict *far-field* pressure distributions and sound power spectra (narrow band and one-third-octave). Validation was performed against reference systems (laser Doppler vibrometer and microphone arrays) to assess the accuracy of both vibration reconstruction and acoustic prediction. To summarize, the work was organised around three complementary case studies. First, the baffled cylindrical shell served as a benchmark geometry admitting analytical treatment of the acoustic field; it was used to verify the end-to-end measurement-to-acoustics chain and to validate the numerical implementation of the Kirchhoff–Helmholtz procedure. Second, a more complex, industrially relevant box-shaped thin-shell structure was investigated to test the methodology on a geometry without *closed-form* solutions: vibration fields measured optically were used to drive a discretised acoustic model, and two approaches for computing sound power were compared: the classical far-field (*ISO-standards*) integration and a novel collapsed-surface implementation that projects the acoustic measurement surface onto the radiating structure. Third, the feasibility of using synchronous event cameras for vibration measurement was demonstrated on a cantilever beam: neural network–based frame reconstruction, marker tracking, and multi-view triangulation were combined and cross-validated against a scanning *LDV*.

9.1 BAFFLED CYLINDER CONCLUSIONS

Chapters 5 and 6 introduced and validated a method for measuring and reconstructing the three-dimensional sound radiation of closed-shell structures using a single high-speed optical camera. The approach was demonstrated on a baffled thin-walled cylindrical structure, with results validated against vibration measurements obtained from a scanning laser vibrometer and acoustic measurements collected with a microphone array. On average, the reconstructed vibration fields deviated by $0.7 - 3.2\ dB$ from the laser vibrometer measurements, while the reconstructed acoustic fields deviated by $0.4 - 2.0\ dB$ from the microphone array measurements. The method also proved suitable for reconstructing narrow-band and one-third-octave spectra of radiated sound power, showing good agreement with reference measurements. In most one-third-octave bands, the mismatch remained within $1 - 2\ dB$, with a slightly larger discrepancy of about $3\ dB$ observed at $630\ Hz$, which may also reflect uncertainties in the reference data. A general formulation was introduced in which the total radiated sound power is derived by integrating the sound intensity over the surface of the structure, approximated through a Riemann mid-point summation over a dense grid of elemental radiators. The combination of camera acquisitions with digital image correlation and photogrammetry enables the reconstruction of the full vibration field of the structure from a relatively small number of short recordings taken from multiple viewpoints. Compared with classical approaches that rely on integrating sound intensity over large measurement surfaces, the proposed method has achieved comparable or higher accuracy. The method offers several advantages: it can be implemented *in situ* without requiring special acoustic facilities such as anechoic or reverberant rooms; the optical setup does not interfere with either the vibration response or the sound radiation; reflections, flanking sources, and background noise have negligible influence in sufficiently large and damped environments; *full-field* optical acquisition enables dense and accurate mapping of both vibration and sound radiation fields; the recording time is relatively short, thus reducing sensitivity to variability and long-term disturbances. At the same time, several limitations must be acknowledged. First, the vibration field is reconstructed in terms of displacements, whose amplitudes decrease with frequency and are smaller than velocities or accelerations, which limits accuracy at higher frequencies. Second, the post-processing is computationally demanding, requiring powerful hardware to reconstruct both vibration and acoustic fields. Third, successful application of *DIC* and photogrammetry requires the radiating surfaces to be prepared with a speckle pattern to ensure accurate point tracking. Finally, the method depends on high-speed, high-resolution cameras to adequately capture wide audio-frequency bands. Overall, the method enhances and simplifies the measurement and reconstruction of structural vibrations and radiated sound power, while showing strong agreement with reference measurements. Although computational effort and surface preparation remain practical challenges, these limitations can be alleviated through advances in optical measurement techniques, post-processing algorithms, and camera technology.

9.2 SMALL THIN-WALLED BOX CONCLUSIONS

Chapters 5 and 7 have extended the investigation from controlled laboratory structures to a more application-oriented case study, focusing on a flexible box representative of machine enclosures. Unlike the cylindrical reference structure, which primarily served as a methodological validation in an idealized laboratory setting, this case emphasizes the applicability of *vision-based* techniques to realistic machine-like systems. The spectra reconstructed from the camera measurements showed overall good agreement with those obtained from direct vibration measurements using a scanning laser vibrometer and from acoustic measurements collected with a microphone array arranged according to *ISO standards*. In particular, the third-octave-band sound power levels derived from the camera acquisitions generally differed by less than $1 - 2\ dB$ from the reference measurements, with somewhat larger discrepancies around the $630\ Hz$ and $1000\ Hz$ bands. These differences, however, are comparable to the uncertainty affecting the reference vibro-acoustic data and do not compromise the overall validity of the method. The novel formulation proposed in this research work for sound power estimation, originally introduced and validated on the cylindrical structure, is here applied to the box, which represents a more realistic machine-like enclosure characterized by structural discontinuities. In this case, the approach has been implemented through a Boundary Element Method framework, which is particularly well suited for handling complex geometries and non-ideal features. By combining *camera-based* vibration reconstructions with *BEM*, the radiated acoustic field can be evaluated directly from the surface data. Compared with the classical *far-field* formulation, this method achieves higher accuracy. The consistency between the *vibration-based* derivation of sound power and the *microphone-based* reference data provides strong evidence of the robustness of this formulation. Again, the method offers several advantages. First, it can be applied in situ without requiring specialized acoustic environments such as anechoic or reverberant rooms, provided that background noise does not significantly affect the structural vibration itself. Second, optical recordings are fast, typically requiring only a few minutes, thereby minimizing sensitivity to long-term disturbances and facilitating the acquisition of representative vibration and sound radiation fields. Moreover, the dense full-field vibration maps obtained through *DIC* photogrammetry enable both a detailed characterization of the structural response and a more accurate reconstruction of the associated sound power radiation. Nonetheless, the limitations already discussed in Section 9.1 also apply here. In particular, the reliance on displacement-based measurements at high frequencies, the need for computationally demanding post-processing, and the requirement of a speckle pattern on the measurement surface remain inherent challenges of the method. Overall, the method enhances and simplifies the measurement and reconstruction of structural vibrations and radiated sound power, providing a robust and versatile framework that complements classical vibro-acoustic techniques and offers a solid basis for further methodological developments and *in situ* applications.

9.3 EVENT CAMERA CONCLUSIONS

As presented in Chapter 8, this part of research investigated the feasibility of reconstructing the flexural vibration field of a cantilever beam using photogrammetry applied to event cameras. Unlike conventional frame-based sensors, event cameras operate asynchronously, capturing pixel-level brightness changes with minimal latency and without redundant data. The proposed methodology was structured around four main components: event data processing, camera calibration, marker tracking and $3D$ triangulation, and vibration analysis. Event data streams were converted into intensity images using the pre-trained neural network *E2VID*, which enabled accurate frame reconstruction even at high sampling frequencies. Calibration was performed through standard photogrammetric procedures, while marker detection and tracking were carried out on rectified frames with subpixel refinement. The resulting trajectories were triangulated using a *Direct Linear Transform approach*, followed by bundle adjustment to refine both camera parameters and marker positions. A key aspect of the experimental setup was the hardware synchronization of up to four event cameras, guaranteeing sub-microsecond temporal alignment. This ensured that the dynamic response could be consistently captured across all views. Two separate tests were conducted, corresponding to the first and second flexural resonance frequencies of the beam (around $32\ Hz$ and around $190\ Hz$). A parametric analysis was carried out to assess the influence of the number of cameras. Starting from four simultaneously acquired views, subsets of two cameras were considered to evaluate the robustness of the triangulation. The comparison against laser Doppler vibrometer reference measurements demonstrated that although satisfactory reconstructions can be achieved with two cameras, the use of four cameras significantly improves robustness by reducing orientation-dependent errors. For instance, considering the first flexural mode, the root mean square error E_{RMS,u_n} between event-based and *LDV* measurements was $0.18\ mm$ for the *1–2* configuration, increased to $0.24\ mm$ for the *1–3* configuration, and resulted in $0.22\ mm$ when all four cameras were employed. This shows that the full-camera setup, although not always yielding the lowest *RMS* error, balances orientation effects and provides redundancy, thereby enhancing the overall reliability of the measurement system. Two additional considerations highlight the potential of this approach. First, the experiments were conducted with relatively low-resolution event cameras; therefore, further improvements in accuracy can be expected with higher-resolution sensors. Second, while this study focused on the first two resonance modes, the methodology is inherently capable of extending to higher frequencies, where the main limiting factor would be the reduced displacement amplitudes rather than the temporal resolution. In summary, the results presented in Chapter 8 demonstrate that event cameras, when coupled with a dedicated photogrammetric post processing, provide a viable and promising alternative for high-speed vibration monitoring. The method combines accuracy, high temporal resolution, and data efficiency, and its scalability from two to four cameras ensures robustness across different viewing geometries. Although computational demands and surface preparation remain practical challenges, these limitations are balanced by the advantages of an optical, non-intrusive, and temporally precise measurement technique, offering significant potential for future applications in structural dynamics and vibro-acoustic analysis. In perspective, the high temporal resolution and efficient data handling of event cameras, recording only pixel-level brightness changes rather than full frames, make them particularly well suited for extending vibro-acoustic analyses beyond laboratory validation. Their ability to capture rapid structural dynamics with reduced storage requirements positions them as promising candidates for applications on more complex radiating systems such as the cylindrical shell and the machine-like box discussed in Chapters 5–7. In such contexts, event-based vision could provide in situ, non-intrusive measurements of both vibration fields and the resulting acoustic radiation, including sound power estimation.

Across these studies the thesis delivers several methodological and practical contributions: demonstration of high-fidelity vibration reconstruction from high-speed and event-based cameras; implementation and validation of Kirchhoff–Helmholtz and *BEM* procedures driven by *full-field* optical data; derivation of narrow band and one-third-octave sound power spectra from optically reconstructed velocities; introduction and assessment of a collapsed-surface approach to sound power estimation; and a systematic experimental quantification of errors (vibration and acoustic) versus reference measurements. Results show that *camera-based* reconstructions typically deviate from *LDV* benchmarks by only a few dB (or sub-millimetre in displacement *RMS* for the event camera experiments), and that the optically driven acoustic predictions reproduce microphone measurements with deviations that are acceptable for many engineering applications. Overall, the three case studies demonstrate that *vision-based* sensing constitutes a practical and broadly applicable methodology for reconstructing structural vibrations and deriving the associated acoustic radiation, including sound power, both under controlled laboratory conditions and, potentially, in real world environments.

PUBLICATIONS

INTERNATIONAL PEER REVIEWED JOURNAL PUBLICATIONS

- Baldini, S., Guernieri, G., Gorjup, D., Gardonio, P., Slavič, J., and Rinaldo, R., 2025, “3D Sound Radiation Reconstruction from Camera Measurements,” Mech Syst Signal Process, 227. https://doi.org/10.1016/j.ymssp.2025.112400.

- Gardonio, P., Baldini, S., Gorjup, D., Slavič, J., and Rinaldo, R., 2025, “Narrow and Third-Octave Spectra of Total Flexural Vibration and Sound Power Radiation from Optical Camera Measurements”, Mech Syst Signal Process, (under review)

- Baldini, S., Deckers, E., Denayer, H., Dong, L., Gardonio, P., Gorjup, D., Panagiotopoulos, D., Rinaldo, R., and Slavič, J., 2025, “The measurement of machinery Sound Power Level (SWL) with optical cameras: theory and experimental validation”, Journal of Sound and Vibration, (under review)

- Baldini, S., Stazi, F., Bernardini, R., Fusiello, A., Gardonio, P., and Rinaldo, R., 2025, “Vibration Measurement with Neuromorphic Vision Sensors”, Mech Syst Signal Process, (under review)

FULL PAPERS IN PROCEEDINGS OF INTERNATIONAL CONFERENCES

- Baldini S., Guernieri G., Gorjup D., Gardonio P., Slavič J., and Rinaldo R., 2024, “Reconstruction of Cylinder Flexural Vibration and Sound Radiation Fields from Multi-View Single-Camera Measurements,” *Proceedings of ISMA 2024*, KU Leuven, Leuven, pp. 2430–2441

- Baldini, S., Bernardini R., Fusiello A., Gardonio P., and Rinaldo R., 2024, “Vibration Measurement with Event Cameras,” *Proceedings of ISMA2024*, Katholieke Universiteit Leuven, Leuven

- Baldini S., Bernardini R., Fusiello A., Gardonio P., and Rinaldo R., 2024, “Measuring Vibrations with Event Cameras,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Brescia. https://doi.org/https://doi.org/10.5194/isprs-archives-XLVIII-2-W7-2024-9-2024.

- Gorjup, D., Baldini, S., Guernieri, G., Gardonio, P., Slavič, J., and Rinaldo, R., 2025, “Frequency-Domain Triangulation for Single-Camera Sound Radiation Measurement,” Proceedings of the 11th International Operational Modal Analysis Conference, IOMAC 2025, Rennes, pp. 185–191

- Baldini, S., Bernardini R., Fusiello A., Gardonio P., and Rinaldo R., 2025, “MATLAB Functions for the Reconstruction of Vibration Fields from Event Camera Acquisitions”, *Proceedings of OpenSD2025,* University of Ljubljana, Ljubljana

- Gardonio P., Baldini S., Gorjup D., Slavič J., and Rinaldo R., 2025, “Sound Power Radiation from Vibration Measurements with Optical Methods: MATLAB functions”, *Proceedings of OpenSD2025,* University of Ljubljana, Ljubljana

REFERENCES

[1] Thomson, W. T., 2017, *Theory of Vibration with Applications*, CRC Press Taylor & Francis Group.

[2] Fahy, F., 2015, *Measurement of Audio-Frequency Sound in Air, in: Fundamentals of Sound and Vibration*, CRC Press.

[3] Fahy, F., and Thompson, D., eds., 2015, *Fundamentals of Sound and Vibration*, CRC Press. https://doi.org/10.1201/b18348.

[4] Fahy, F., and Walker, J., 2004, *Advanced Applications in Acoustics, Noise and Vibration*, CRC Press.

[5] Kenneth G. McConnell, and Paulo S. Varoto, 2008, *Vibration Testing: Theory and Practice*, New York.

[6] Austin R. Frey, Alan B. Coppens, and James V. Sanders, 2000, *Fundamentals of Acoustics*, New York.

[7] Verheij, J. W., 1997, "Inverse and Reciprocity Methods for Machinery Noise Source Characterization and Sound Path Quantification Part 1: Sources," The International Journal of Acoustics and Vibration, 2(1). https://doi.org/10.20855/ijav.1997.2.107.

[8] D. J. Ewins, 2009, *Modal Testing: Theory, Practice and Application*, Research Studies Pre.

[9] Barron, R. F., 2003, *Industrial Noise Control and Acoustics*, Marcel Dekker Inc., New York.

[10] Bies, D. A., and Hansen, C. H., 2009, *Engineering Noise Control Theory and Practice*, Spon Press, London.

[11] Cuyx, B., Desmet, W., Buyens, W., and Waterschoot, T. van, 2020, "Design and Validation of a Low-Cost Acoustic Anechoic Chamber," Journal of the Audio Engineering Society, 149, pp. 1–5.

[12] Habault, D., Friot, E., Herzog, P., and Pinhede, C., 2017, "Active Control in an Anechoic Room: Theory and First Simulations," Acta acustica united with Acustica, 103(3), pp. 369–378. https://doi.org/10.3813/aaa.919066.

[13] Pinhède, C., Habault, D., Friot, E., and Herzog, Ph., 2021, "Active Control of the Field Scattered by the Rigid Wall of a Semi-Anechoic Room—Simulations and Full-Scale off-

Line Experiment," J Sound Vib, 506, p. 116134. https://doi.org/10.1016/j.jsv.2021.116134.

[14] Pinhède, C., Boulandet, R., Friot, E., Allado, M., Côte, R., and Herzog, P., 2022, "Towards an Active Semi-Anechoic Room : Simulations and First Measurements," *Proceedings of the 10th Convention of the European Acoustics Association Forum Acusticum 2023*, European Acoustics Association, Turin, Italy, pp. 3611–3618. https://doi.org/10.61782/fa.2023.0399.

[15] Michel, U., 2006, "History of Acoustic Beamforming," *Proceedings of the 1st Berlin Beamforming Conference*, pp. 22–23.

[16] Merino-Martínez, R., Sijtsma, P., Snellen, M., and others, 2019, "A Review of Acoustic Imaging Methods Using Phased Microphone Arrays," CEAS Aeronaut. J., 48, pp. 197–230.

[17] Boone, M. M., Kinneging, N., and van den Dool, T., 2000, "Two-Dimensional Noise Source Imaging with a T-Shaped Microphone Cross Array," J Acoust Soc Am, 108(6), pp. 2884–2890. https://doi.org/10.1121/1.1320477.

[18] Boone, M. M., and Berkhout, A. J., 1984, "Theory and Applications of a High-Resolution Synthetic Acoustic Antenna for Industrial Noise Measurements," Noise Control Eng J, 23(2), p. 60. https://doi.org/10.3397/1.2827639.

[19] BRÜHL, S., and RÖDER, A., 2000, "ACOUSTIC NOISE SOURCE MODELLING BASED ON MICROPHONE ARRAY MEASUREMENTS," J Sound Vib, 231(3), pp. 611–617. https://doi.org/10.1006/jsvi.1999.2548.

[20] Mellet, C., Létourneaux, F., Poisson, F., and Talotte, C., 2006, "High Speed Train Noise Emission: Latest Investigation of the Aerodynamic/Rolling Noise Contribution," J Sound Vib, 293(3–5), pp. 535–546. https://doi.org/10.1016/j.jsv.2005.08.069.

[21] Nordborg, A., Wedemann, J., and Willenbrink, L., 2000, "Optimum Array Microphone Configuration," *Proceedings Inter-Noise*, Nice, France.

[22] Billingsley, J., and Kinns, R., 1976, "The Acoustic Telescope," J Sound Vib, 48(4), pp. 485–510. https://doi.org/10.1016/0022-460x(76)90552-6.

[23] Ning, F., Meng, D., and Wei, J., 2022, "An Improved Acoustic Imaging Algorithm Combining Object Detection and Beamforming for Acoustic Camera," JASA Express Lett, 2(6). https://doi.org/10.1121/10.0011735.

[24] Holland, K. R., and Nelson, P. A., 2012, "An Experimental Comparison of the Focused Beamformer and the Inverse Method for the Characterisation of Acoustic Sources in Ideal and Non-Ideal Acoustic Environments," J Sound Vib, 331(20), pp. 4425–4437. https://doi.org/10.1016/j.jsv.2012.05.005.

[25] Brooks, T. F., and Humphreys, W. M., 2006, "A Deconvolution Approach for the Mapping of Acoustic Sources (DAMAS) Determined from Phased Microphone Arrays," J Sound Vib, 294(4–5), pp. 856–879. https://doi.org/10.1016/j.jsv.2005.12.046.

[26] Fahy, F., 1995, *Sound Intensity*, CRC Press.

[27] Jacobsen, F., and Bree, H.-E. de, 2005, "A Comparison of Two Different Sound Intensity Measurement Principles," J Acoust Soc Am, 118(3), pp. 1510–1517. https://doi.org/10.1121/1.1984860.

[28] Fahy, F. J., 1977, "Measurement of Acoustic Intensity Using the Cross-Spectral Density of Two Microphone Signals," J Acoust Soc Am, 62(4), pp. 1057–1059. https://doi.org/10.1121/1.381601.

[29] Maynard, J. D., Williams, E. G., and Lee, Y., 1985, "Nearfield Acoustic Holography: I. Theory of Generalized Holography and the Development of NAH," J Acoust Soc Am, 78(4), pp. 1395–1413. https://doi.org/10.1121/1.392911.

[30] Williams, E. G., Maynard, J. D., and Skudrzyk, E., 1980, "Sound Source Reconstructions Using a Microphone Array," J Acoust Soc Am, 68(1), pp. 340–344. https://doi.org/10.1121/1.384602.

[31] Hald, J., *Time Domain Acoustical Holography and Its Applications*.

[32] Geng, L., Chen, X.-G., He, C.-D., Chen, W., and He, S.-P., 2023, "Compressive Nonstationary Near-Field Acoustic Holography for Reconstructing the Instantaneous Sound Field," Mech Syst Signal Process, 204, p. 110779. https://doi.org/10.1016/j.ymssp.2023.110779.

[33] Salin, M. B., and Kosteev, D. A., 2020, "Nearfield Acoustic Holography-Based Methods for Far Field Prediction," Applied Acoustics, 159, p. 107099. https://doi.org/10.1016/j.apacoust.2019.107099.

[34] Geng, L., Chen, X.-G., He, S.-P., and He, C.-D., 2024, "Reconstruction of Transient Acoustic Field Using Sparse Real-Time near-Field Acoustic Holography," J Sound Vib, 568, p. 117973. https://doi.org/10.1016/j.jsv.2023.117973.

[35] Wu, H., Zha, Y., Gao, J., Zhou, H., Zhang, N., and Jiang, W., 2024, "A Hybrid Nearfield Acoustic Holography Based on the Mapping Relationship and Boundary Element Method," Applied acoustics, 216, p. 109823. https://doi.org/10.1016/j.apacoust.2023.109823.

[36] R. Burtch, 2004, *History of Photogrammetry*.

[37] Gruen, A., 1997, "Fundamentals of Videogrammetry — A Review," Hum Mov Sci, 16(2–3), pp. 155–187. https://doi.org/10.1016/S0167-9457(96)00048-6.

[38] Yu, L., and Pan, B., 2017, “Single-Camera High-Speed Stereo-Digital Image Correlation for Full-Field Vibration Measurement,” Mech. Syst. Signal Pr., 94, pp. 374–383.

[39] Gorjup, D., Slavi\u{c}, J., and Boltezar, M., 2019, “Frequency Domain Triangulation for Full-Field 3D-Deflection-Shape Identification,” Mech. Syst. Signal Pr., 133 106287, pp. 143–152.

[40] Gorjup, D., Slavi\u{c}, J., Babnik, A., and Boltezar, M., 2021, “Still-Camera Multiview Spectral Optical Flow Imaging for 3D Operating-Deflection-Shape Identification,” Mech. Syst. Signal Pr., 152 107456.

[41] Luo, P. F., Chao, Y. J., Sutton, M. A., and Peters, W. H., 1993, “Accurate Measurement of Three-Dimensional Deformations in Deformable and Rigid Bodies Using Computer Vision,” Exp Mech, 33(2), pp. 123–132. https://doi.org/10.1007/BF02322488.

[42] Peters, W. H., and Ranson, W. F., 1982, “Digital Imaging Techniques In Experimental Stress Analysis,” Optical Engineering, 21(3). https://doi.org/10.1117/12.7972925.

[43] Chu, T. C., Ranson, W. F., and Sutton, M. A., 1985, “Applications of Digital-Image-Correlation Techniques to Experimental Mechanics,” Exp Mech, 25(3), pp. 232–244. https://doi.org/10.1007/BF02325092.

[44] Sutton, M., Wolters, W., Peters, W., Ranson, W., and McNeill, S., 1983, “Determination of Displacements Using an Improved Digital Correlation Method,” Image Vis Comput, 1(3), pp. 133–139. https://doi.org/10.1016/0262-8856(83)90064-1.

[45] Hubert Schreier, Jean-José Orteu, and Michael A. Sutton, *Image Correlation for Shape, Motion and Deformation Measurements*, Springer.

[46] Liu, T., Burner, A. W., Jones, T. W., and Barrows, D. A., 2012, “Photogrammetric Techniques for Aerospace Applications,” Progress in Aerospace Sciences, 54, pp. 1–58. https://doi.org/10.1016/j.paerosci.2012.03.002.

[47] Su, X., and Zhang, Q., 2010, “Dynamic 3-D Shape Measurement Method: A Review,” Opt Lasers Eng, 48(2), pp. 191–204. https://doi.org/10.1016/j.optlaseng.2009.03.012.

[48] Jiang, R., Jáuregui, D. V., and White, K. R., 2008, “Close-Range Photogrammetry Applications in Bridge Measurement: Literature Review,” Measurement, 41(8), pp. 823–834. https://doi.org/10.1016/j.measurement.2007.12.005.

[49] Javh, J., Slavič, J., and Boltežar, M., 2017, “The Subpixel Resolution of Optical-Flow-Based Modal Analysis,” Mech Syst Signal Process, 88, pp. 89–99. https://doi.org/10.1016/j.ymssp.2016.11.009.

[50] Baldini, S., Guernieri, G., Gorjup, D., Gardonio, P., Slavič, J., and Rinaldo, R., 2025, "3D Sound Radiation Reconstruction from Camera Measurements," Mech Syst Signal Process, 227. https://doi.org/10.1016/j.ymssp.2025.112400.

[51] Cheng, Y., Tian, Z., Ning, D., Feng, K., Li, Z., Chauhan, S., and Vashishtha, G., 2025, "Computer Vision-Based Non-Contact Structural Vibration Measurement: Methods, Challenges and Opportunities," Measurement, 243, p. 116426. https://doi.org/10.1016/j.measurement.2024.116426.

[52] ISO (International Organization for Standardization), 2012, "ISO 3745:2012: Acoustics — Determination of Sound Power Levels and Sound Energy Levels of Noise Sources Using Sound Pressure — Precision Methods for Anechoic Rooms and Hemi-Anechoic Rooms."

[53] ISO (International Organization for Standardization), 2010, "ISO 3741:2010: Acoustics — Determination of Sound Power Levels and Sound Energy Levels of Noise Sources Using Sound Pressure — Precision Methods for Reverberation Test Rooms."

[54] ISO (International Organization for Standardization), 2010, "ISO1 3744:2010: Acoustics - Determination of Sound Power Levels and Sound Energy Levels of Noise Sources Using Sound Pressure - Engineering Methods for an Essentially Free Field over a Reflecting Plane ."

[55] Pierce, A. D., 1991, *Acoustics*, Acoustical Society of America through the American Institute of Physics.

[56] Junger, M. C., and Feit, D., 2010, *Sound, Structures, and Their Interaction*, The Mit Press, Imp.

[57] Lesuer, C., and Gotteland, M., 1988, "Rayonnement Acoustique. Éléments de Base," *Rayonnement Acoustique Des Structures*, y C. Lesueur Editions Eyrolles, p. Ch 3.

[58] Fahy, F. J., and Gardonio, P., 2007, *Sound and Structural Vibration*, Elsevier.

[59] Rohlfing, J., Gardonio, P., and Thompson, D. J., 2011, "Comparison of Decentralized Velocity Feedback Control for Thin Homogeneous and Stiff Sandwich Panels Using Electrodynamic Proof-Mass Actuators," J Sound Vib, 330(5), pp. 843–867. https://doi.org/10.1016/j.jsv.2010.09.013.

[60] Elliott, S. J., and Johnson, M. E., 1993, "Radiation Modes and the Active Control of Sound Power," J Acoust Soc Am, 94(4), pp. 2194–2204. https://doi.org/10.1121/1.407490.

[61] Jones, C. B., Goates, C. B., Blotter, J. D., and Sommerfeldt, S. D., 2020, "Experimental Validation of Determining Sound Power Using Acoustic Radiation Modes and a Laser Vibrometer," Applied Acoustics, 164, p. 107254. https://doi.org/10.1016/j.apacoust.2020.107254.

[62] Bacon, I. C., Sommerfeldt, S. D., and Blotter, J. D., 2022, "Developing an Indirect Vibration-Based Sound Power Method to Determine the Sound Power Radiated from Acoustic Sources," p. 065003. https://doi.org/10.1121/2.0001732.

[63] Baqersad, J., Poozesh, P., Niezrecki, C., and Avitabile, P., 2017, "Photogrammetry and Optical Methods in Structural Dynamics – A Review," Mech Syst Signal Process, 86, pp. 17–34. https://doi.org/10.1016/j.ymssp.2016.02.011.

[64] Gardonio, P., Rinaldo, R., Dal Bo, L., Del Sal, R., Turco, E., and Fusiello, A., 2022, "Free-Field Sound Radiation Measurement with Multiple Synchronous Cameras," Measurement, 188, p. 110605. https://doi.org/10.1016/j.measurement.2021.110605.

[65] Gallego, G., Delbrück, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A. J., Conradt, J., Daniilidis, K., and Scaramuzza, D., 2022, "Event-Based Vision: A Survey," IEEE Trans Pattern Anal Mach Intell, 44(1), pp. 154–180. https://doi.org/10.1109/TPAMI.2020.3008413.

[66] iniVation, A. G., 2020, "Understanding the Performance of Neuromorphic Event-Based Vision," Tech. Rep.

[67] Amir, A., Taba, B., Berg, D., Melano, T., McKinstry, J., Di Nolfo, C., Nayak, T., Andreopoulos, A., Garreau, G., Mendoza, M., Kusnitz, J., Debole, M., Esser, S., Delbruck, T., Flickner, M., and Modha, D., 2017, "A Low Power, Fully Event-Based Gesture Recognition System," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp. 7388–7397. https://doi.org/10.1109/CVPR.2017.781.

[68] Baldini, S., Bernardini R., Fusiello A., Gardonio P., and Rinaldo R., 2024, "Vibration Measurement with Event Cameras," *Proceedings of ISMA2024*, Katholieke Universiteit Leuven, Leuven.

[69] Sofia Baldini, Riccardo Bernardini, Andrea Fusiello, Paolo Gardonio, and Roberto Rinaldo, 2024, "Measuring Vibrations with Event Cameras," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Brescia. https://doi.org/https://doi.org/10.5194/isprs-archives-XLVIII-2-W7-2024-9-2024.

[70] Fusiello, A., 2024, *Computer Vision: Three-Dimensional Reconstruction Techniques*, Springer Cham.

[71] Richard Hartley, and Andrew Zisserman, 2004, *Multiple View Geometry in Computer Vision*, Cambridge University Press.

[72] Woodham, R. J., 1980, "Photometric Method For Determining Surface Orientation From Multiple Images," Optical Engineering, 19(1). https://doi.org/10.1117/12.7972479.

[73] Kuroda, T., 2017, *Essential Principles of Image Sensors*, CRC Press. https://doi.org/10.1201/b17411.

[74] Kihong Shin, and Joseph Hammond, 2008, *Fundamentals of Signal Processing for Sound and Vibration Engineers*, Wiley.

[75] Boahen, K. A., 2000, "Point-to-Point Connectivity between Neuromorphic Chips Using Address Events," IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 47(5), pp. 416–434. https://doi.org/10.1109/82.842110.

[76] Purohit, P., and Manohar, R., 2022, "Field-Programmable Encoding for Address-Event Representation," Front Neurosci, 16. https://doi.org/10.3389/fnins.2022.1018166.

[77] Sturm, P. F., and Maybank, S. J., 1999, "On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications," *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, pp. 432–437. https://doi.org/10.1109/CVPR.1999.786974.

[78] Zhang, Z., 2000, "A Flexible New Technique for Camera Calibration," IEEE Trans Pattern Anal Mach Intell, 22(11), pp. 1330–1334. https://doi.org/10.1109/34.888718.

[79] Chu, T. C., Ranson, W. F., and Sutton, M. A., 1985, "Applications of Digital-Image-Correlation Techniques to Experimental Mechanics," Exp Mech, 25(3), pp. 232–244. https://doi.org/10.1007/BF02325092.

[80] Pan, B., 2018, "Digital Image Correlation for Surface Deformation Measurement: Historical Developments, Recent Advances and Future Goals," Meas Sci Technol, 29(8), p. 082001. https://doi.org/10.1088/1361-6501/aac55b.

[81] Kuhn, H. W., 1955, "The Hungarian Method for the Assignment Problem," Naval Res. Logist. Quart., 2(1–2), pp. 83–97.

[82] Kahn-Jetter, Z. L., and Chu, T. C., 1990, "Three-Dimensional Displacement Measurements Using Digital Image Correlation and Photogrammic Analysis," Exp Mech, 30(1), pp. 10–16. https://doi.org/10.1007/BF02322695.

[83] M. A. Sutton, J. J. Orteu, and H. Schreier, 2009, *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications*, Springer Science & Business Media, New York.

[84] Morse, P. M., and Ingard, K. U., 1986, *Theoretical Acoustics*, Princeton University Press.

[85] Gaul, L., Kogl, M., and Wagner, M., 2013, *Boundary Element Methods for Engineers and Scientists*, Springer Science & Business Media. [Online]. Available: http://www.springer.de/engine/.

[86] Baldini Sofia, Guernieri Gianluca, Gorjup Domen, Gardonio Paolo, Slavič Janko, and Rinaldo Roberto, 2024, “Reconstruction of Cylinder Flexural Vibration and Sound Radiation Fields from Multi-View Single-Camera Measurements,” *Proceedings of ISMA 2024*, KU Leuven, Leuven, pp. 2430–2441.

[87] Gorjup, D., Baldini, S., Guernieri, G., Gardonio, P., Slavič, J., and Rinaldo, R., 2025, “Frequency-Domain Triangulation for Single-Camera Sound Radiation Measurement,” Proceedings of the 11th International Operational Modal Analysis Conference, IOMAC 2025, Rennes, pp. 185–191.

[88] MARBURG, S., 2002, “SIX BOUNDARY ELEMENTS PER WAVELENGTH: IS THAT ENOUGH?,” Journal of Computational Acoustics, 10(01), pp. 25–51. https://doi.org/10.1142/S0218396X02001401.

[89] Del Sal, R., Dal Bo, L., Turco, E., Fusiello, A., A. Zanarini, Rinaldo, R., and Gardonio, P., 2021, “Structural Vibration Measurement with Multiple Synchronous Cameras,” Mech. Syst. Signal Pr., 157, p. 107742.

[90] Lucas, B., and Kanade, T., 1981, *An Iterative Image Registration Technique with an Application to Stereo Vision (IJCAI)*.

[91] Werner Soedel, 2004, *Vibrations of Shells and Plates*, CRC Press.

[92] Robert D. Blevins, 1979, *Formulas for Natural Frequency and Mode Shape*, New York: Van Nostrand Reinhold Co.

[93] Reddy, J. N., 2007, *Theory and Analysis of Elastic Plates and Shells*, CRC Press.

[94] Leissa, A. W., 1973, *Vibration of Shells*, Washington.

[95] Stephan P. Timoshenko, 1964, *Theory of Plates and Shells*.

[96] Leissa A. W., 1969, *Vibration of Plates*, Washington.

[97] Gardonio, P., and Turco, E., 2019, “Tuning of Vibration Absorbers and Helmholtz Resonators Based on Modal Density/Overlap Parameters of Distributed Mechanical and Acoustic Systems,” J Sound Vib, 451, pp. 32–70. https://doi.org/10.1016/j.jsv.2019.03.015.

[98] Markuš, Š., 1998, *The Mechanics of Vibrations of Cylindrical Shells*, Elsevier.

[99] Blevins, R. D., 2016, *Formulas for Dynamics, Acoustics and Vibration*, John Wiley & Sons.

[100] Koopmann, G. H., and Fahnline, J. B., 1997, *Designing Quiet Structures. A Sound Power Minimization Approach*, Academic Press.

[101] Stepanishen, P. R., 1978, “Radiated Power and Radiation Loading of Cylindrical Surfaces with Nonuniform Velocity Distributions,” J Acoust Soc Am, 63(2), pp. 328–338. https://doi.org/10.1121/1.381743.

[102] Sun, Y., Yang, T., and Chen, Y., 2018, “Sound Radiation Modes of Cylindrical Surfaces and Their Application to Vibro-Acoustics Analysis of Cylindrical Shells,” J Sound Vib, 424, pp. 64–77. https://doi.org/10.1016/j.jsv.2018.03.004.

[103] Goates, C. B., Jones, C. B., Sommerfeldt, S. D., and Blotter, J. D., 2020, “Sound Power of Vibrating Cylinders Using the Radiation Resistance Matrix and a Laser Vibrometer,” J Acoust Soc Am, 148(6), pp. 3553–3561. https://doi.org/10.1121/10.0002870.

[104] Skudrzyk, E., 1971, *The Foundations of Acoustics*, Springer-Verlag.

[105] Heisterkamp, F., Schmitt, J., and Hook, J., 2023, “Simplification of ISO 3744 - Making the Most Important Standard for Determination of Sound Power Levels Easier to Use,” INTER-NOISE and NOISE-CON Congress and Conference Proceedings, 265(7), pp. 319–326. https://doi.org/10.3397/IN_2022_0049.

[106] Milton, J., Cheer, J., and Daley, S., 2019, “Experimental Identification of the Radiation Resistance Matrix,” J Acoust Soc Am, 145(5), pp. 2885–2894. https://doi.org/10.1121/1.5102167.

[107] Vicente Cutanda Henríquez, and Peter Møller Juhl, 2010, “OpenBEM-an Open Source Boundary Element Method Software in Acoustics,” *Internoise 2010*, Sociedade Portuguesa de Acustica (SPA), Lisbon.

[108] Geuzaine, C., and Remacle, J., 2009, “Gmsh: A 3-D Finite Element Mesh Generator with Built-in Pre- and Post-processing Facilities,” Int J Numer Methods Eng, 79(11), pp. 1309–1331. https://doi.org/10.1002/nme.2579.

[109] Marburg, S., “Discretization Requirements: How Many Elements per Wavelength Are Necessary?,” *Computational Acoustics of Noise Propagation in Fluids - Finite and Boundary Element Methods*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 309–332. https://doi.org/10.1007/978-3-540-77448-8_12.

[110] Steffen Marburg, and Bodo Nolte, eds., 2008, *Computational Acoustics of Noise Propagation in Fluids - Finite and Boundary Element Methods*, Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-77448-8.

[111] Conradt, J., Cook, M., Berner, R., Lichtsteiner, P., J. Douglas, R., and Delbruck, T., 2009, “A Pencil Balancing Robot Using a Pair of AER Dynamic Vision,” *IEEE Int. Symp. Circuits Syst. (ISCAS)*, pp. 781–784.

[112] Cook, M., Gugelmann, L., Jug, F., Krautz, C., and Steger, A., 2011, “Interacting Maps for Fast Visual Interpretation,” *Int. Joint Conf. Neural Netw. (IJCNN)*, pp. 770–776.

[113] Benosman, R., Clercq, C., Lagorce, X., Ieng, S.-H., and C. Bartolozzi, 2014, "Event-Based Visual Flow," IEEE Trans. Neural Netw. Learn. Syst., 25(2), pp. 407–417.

[114] Gallego, G., Lund, J. E. A., Mueggler, E., Rebecq, H., T. Delbruck, and Scaramuzza, D., 2018, "Event-Based, 6-DOF Camera Tracking from Photometric Depth Maps," IEEE Trans. Pattern Anal. Mach. Intell., 40(10), pp. 2402–2412.

[115] Lagorce, X., Orchard, G., Gallupi, F., Shi, B. E., and R. Benosman, 2017, "HOTS: A Hierarchy of Event-Based Time-Surfaces for Pattern," IEEE Trans. Pattern Anal. Mach. Intell., 39(7), pp. 1346–1359.

[116] Sironi, A., Brambilla, M., Bourdis, N., Lagorce, X., and R. Benosman, 2018, "HATS: Histograms of Averaged Time Surfaces for Robust Event-Based Classification," *IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, pp. 1731–1740.

[117] Zhou, Y., Gallego, G., Rebecq, H., Kneip, L., Li, H., and D. Scaramuzza, 2018, "Semi-Dense 3D Reconstruction with a Stereo Event Camera," *Eur. Conf. Comput. Vis. (ECCV)*, pp. 242–258.

[118] Ronneberger, O., Fischer, P., and Brox, T., 2015, "U-Net: Convolutional Networks for Biomedical Image Segmentation," pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

[119] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo, 2015, "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting," *Advances in Neural Information Processing Systems*.

[120] Sutherland, I. E., 1974, "Three-Dimensional Data Input by Tablet," Proceedings of the IEEE, 62(4), pp. 453–461. https://doi.org/10.1109/PROC.1974.9449.

[121] Rebecq, H., Ranftl, R., Koltun, V., and Scaramuzza, D., 2019, "Events-To-Video: Bringing Modern Computer Vision to Event Cameras," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp. 3852–3861. https://doi.org/10.1109/CVPR.2019.00398.