Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Physics

Modeling band structure properties in sonic crystals using machine learning

Master's Thesis

Bc. Aneta Furmanová

Supervisor: RNDr. MgA. Viktor Hruška, Ph.D.

Study Programme: Medical Electronics and Bioinformatics

Field of Study: Bioinformatics

Prague, May 20, 2025

MASTER‘S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Furmanová Aneta** Personal ID number: **499125**
Faculty / Institute: **Faculty of Electrical Engineering**
Department / Institute: **Department of Computer Science**
Study program: **Medical Electronics and Bioinformatics**
Specialisation: **Bioinformatics**

II. Master’s thesis details

Master’s thesis title in English:

Modeling band structure properties in sonic crystals using machine learning

Master’s thesis title in Czech:

Modelovaní vlastností pásové struktury sonických krystalů pomocí strojového učení

Name and workplace of master’s thesis supervisor:

RNDr. MgA. Viktor Hruška, Ph.D. Department of Physics FEE

Name and workplace of second master’s thesis supervisor or consultant:

Date of master’s thesis assignment: **10.02.2025** Deadline for master's thesis submission: **23.05.2025**

Assignment valid until: **20.09.2026**

Head of department’s signature

prof. Mgr. Petr Páta, Ph.D.
Vice-dean´s signature on behalf of the Dean

III. Assignment receipt

The student acknowledges that the master’s thesis is an individual work.
The student must produce her thesis without the assistance of others, with the exception of provided consultations.
Within the master’s thesis, the author must state the names of consultants and include a list of references.

Date of assignment receipt

Student’s signature

MASTER'S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Furmanová Aneta** Personal ID number: **499125**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Computer Science**

Study program: **Medical Electronics and Bioinformatics**

Specialisation: **Bioinformatics**

II. Master's thesis details

Master's thesis title in English:

Modeling band structure properties in sonic crystals using machine learning

Master's thesis title in Czech:

Modelovaní vlastností pásové struktury sonických krystalů pomocí strojového učení

Guidelines:

- Use the review of literature and methods prepared from the diploma project
- Implement numerical solutions of wave equations with non-constant coefficients and periodic boundary conditions of the Floquet-Bloch type
- Analyze the properties of the training dataset (in particular, that they satisfy symmetries based on physical assumptions)
- Obtain analytical formulas describing the band structure properties of sonic crystals using machine learning
- Verify that the discovered formulas are physically interpretable and reasonable (eventually change cost function properties to enforce physics-informed aspects)
- Check that results are consistent with direct FEM simulations
- Support with specific examples from biomedical applications (e.g. consider ultrasound for imaging or traffic noise barriers)

Bibliography / sources:

1. Jiménez, Noé, Olga Umnova, and Jean-Philippe Groby. Acoustic Waves in Periodic Structures, Metamaterials, and Porous Media: From Fundamentals to Industrial Applications. Springer Nature, 2021.
2. Blackstock, David T. Fundamentals of Physical Acoustics. John Wiley & Sons, 2000.
3. Brunton, Steven L., and J. Nathan Kutz. Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. Cambridge University Press, 2022.
4. Camps-Valls, Gustau, Andreas Gerhardus, Urmi Ninad, Gherardo Varando, Georg Martius, Emili Balaguer-Ballester, Ricardo Vinuesa, Emiliano Diaz, Laure Zanna, and Jakob Runge. “Discovering Causal Relations and Equations From Data.” Physics Reports 1044 (November 7, 2023): 1–68. https://doi.org/10.1016/j.physrep.2023.10.005.
5. Gupta, Arpan. “A Review on Sonic Crystal, Its Applications and Numerical Analysis Techniques.” Acoustical Physics 60, no. 2 (March 1, 2014): 223–34. https://doi.org/10.1134/s1063771014020080.

FAKULTA ELEKTROTECHNICKÁ
FACULTY OF ELECTRICAL ENGINEERING
Technická 2
166 27 Praha 6

DECLARATION

I, the undersigned

Student's surname, given name(s):	Furmanová Aneta
Personal number:	499125
Programme name:	Medical Electronics and Bioinformatics

declare that I have elaborated the master’s thesis entitled

Modeling band structure properties in sonic crystals using machine learning

independently, and have cited all information sources used in accordance with the Methodological Instruction on the Observance of Ethical Principles in the Preparation of University Theses and with the Framework Rules for the Use of Artificial Intelligence at CTU for Academic and Pedagogical Purposes in Bachelor’s and Continuing Master’s Programmes.

I declare that I used artificial intelligence tools during the preparation and writing of this thesis. I verified the generated content. I hereby confirm that I am aware of the fact that I am fully responsible for the contents of the thesis.

In Prague on 15.05.2025

Bc. Aneta Furmanová
....
student's signature

Acknowledgements

I would like to express my gratitude to my supervisor, Viktor Hruška, for his guidance, support and encouragement. I am grateful for the opportunities I was given. I would also like to thank my family for their support.

Abstract

This thesis explores an acoustic analogy of electronic band structure. While a wide range of analytical approaches and numerical methods exist to solve for the transmission properties of sonic crystals, existing analytical expressions relating the waveguide geometry to acoustic transmission are limited to a few special cases. The goal is to model the band structure properties of sonic crystals using machine learning since we have access to a numerical model, which presents a valuable opportunity for data-driven discovery. A dataset of numerical solutions was generated using the Webster equation for the unit cell with periodic boundary conditions of the Floquet-Bloch type. To reduce the complexity of the problem, this dataset is transformed into a lower-dimensional space using Principal Component Analysis. In this new coordinate system, lower-dimensional patterns are extracted via symbolic regression implemented in an open-source library PySR. The resulting model is interpretable in terms of underlying physics and can be used, for example, to propose optimized designs for a desired bandgap width for sound barriers. These results enhance the overall understanding of the system, enabling a deeper insight into the underlying principles and providing a more efficient alternative to computationally demanding numerical optimization.

Keywords: locally periodic structures, sonic crystals, machine learning, symbolic regression

Abstrakt

Tato práce se zabývá akustickou analogií elektronové pásové struktury. I když pro řešení přenosových vlastností zvukových krystalů existuje celá řada analytických přístupů a numerických metod, stávající analytické výrazy vztahující geometrii vlnovodu k jeho akustickému přenosu jsou omezeny na několik speciálních případů. Cílem této práce je modelování vlastností pásové struktury sonických krystalů pomocí strojového učení, protože přístup k numerickému modelu představuje cennou příležitost pro nalézání vnitřních zákonitostí na základě analýzy dat. Dataset numerických řešení byl vytvořen pomocí Websterovy rovnice pro jednotkovou buňku s periodickými okrajovými podmínkami Floquetova-Blochova typu. Za účelem snížení složitosti problému je tento dataset transformován do méně rozměrného prostoru pomocí analýzy hlavních komponent. V tomto novém souřadnicovém systému jsou méně dimenzionální vzorce extrahovány pomocí symbolické regrese implementované v knihovně s otevřeným zdrojovým kódem PySR. Výsledný model je interpretovatelný z hlediska vnitřních zákonitostí a lze jej použít například k optimalizaci návrhů pro požadovanou šířku pásma pro hlukové bariéry. Výsledky diplomové práce přispívají k celkovému pochopení systému, přinášejí hlubší vhled do základních principů sonických krystalů a představují efektivnější alternativu k výpočetně náročné numerické optimalizaci.

Klíčová slova: lokálně periodické struktury sonické krystaly, strojové učení, symbolická regrese

Contents

List of Figures

List of Abbreviations

CMA-ES	Covariance Matrix Adaptation Evolution Strategy
FEM	Finite Element Method
GD	Gradient Descent
LASSO	Least Absolute Shrinkage and Selection Operator
LM	Levenberg-Marquardt algorithm
LNSE	Linearized Navier-Stokes Equations
PCA	Principal Component Analysis
RK45	Runge-Kutta-Fehlberg method
SR	Symbolic Regression
TMM	Transfer Matrix Method

Chapter 1

Introduction

The thesis deals with modeling the band structure properties in sonic crystals using machine learning, and this chapter introduces the reader to the topic of periodic structures in acoustics and their applications, such as ultrasound for imaging or noise reduction. Further, it provides an overview of the current state of the art for modeling acoustic transmission in periodic structures and discusses optimization techniques that are currently being used. Finally, the motivation behind the thesis is presented, along with the specific goals.

1.1 Origins of periodic structures

The story of periodic structures for macroscopic engineering applications originates from photonics: in the late 80's, both John [1] and Yablonovitch [2] simultaneously but independently pointed out the propagation properties of electromagnetic waves inside periodic systems. In frequency ranges specific to the configuration, there was no wave propagation in the periodic systems. The frequency ranges were named band gaps by analogy to solid state physics nomenclature. Similarly, their energy ranges where no electronic states exist are called band gaps or energy gaps. Due to their ability to modify propagation the propagation of electromagnetic waves, these so-called photonic crystals have several applications, such as generation, waveguiding, focusing, or splitting of light [3].

A few years later, theoretical works on wave propagation in periodic media emerged also in acoustics, showing the existence of acoustic and elastic bandgaps [4, 5]. Here, the bandgaps present frequency regions, where the wave propagation is significantly reduced. Similarly to photonic crystals, the periodic structures attenuating acoustic and elastic waves were named phononic crystals.

The first constructed example of a phononic crystal is the artwork "Órgano" by sculptor Eusebio Sempere, although he was propably not aware of that. For the sculpture, see Figure 1.1. In 1995, the transmission characteristics by this sculpture was measured by Martínez-Sala [6] and the sound attenuation by sculpture was proven, leading to the first experimental evidence of the presence of acoustic band gaps.

Due to their transmission and reflection characteristics, not achievable by any natural medium [7], the phononic crystals have attracted further attention. The rich publication history is summarized in recent reviews [8, 9].

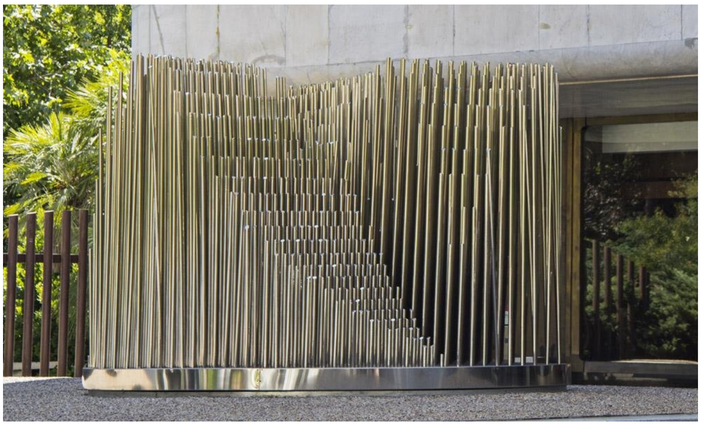


Figure 1.1: Sculpture Órgano by Eusebio Sempere exhibited at the Juan March Foundation in Madrid (from [10]), representing a two-dimensional phononic crystal.

1.2 Phononic and sonic crystals

Before delving into details on bandgaps and specifics of phononic crystals, just a short remark on their actual size and performance. The significant difference between photonic and phononic crystals lies in their size. The periodic structures interact with waves if the dimension of the scatterer and their spacing is of the order of wave length of propagating wave. Further, the bandgap is obtained for an infinite structure (or at least sufficiently large) [7]. Since the photonic crystals can be constructed from scaterrers in the size of microns and it is feasible to do so with thousands of repeating units, no propagation of electromagnetic waves in the bandgap can be guaranteed. On the other hand, the phononic crystals are supposed to attenuate waves in the audible region (i.e., 20 Hz to 20 kHz), which corresponds to wavelengths of meters to centimeters. It is not hard to imagine, that the construction of a large enough structure is spatially demanding. Therefore only locally periodic structures are constructed with significant wave attenuation in the bandgaps.

By altering the way how the phononic crystals are arranged, the occurance of bandgaps can be influenced together with the amount of the wave attenuation. Examples of phononic crystal configurations can be seen in Figure 1.2. The dimensionality of the crystal reflects the arrangement of the scaterrers [11]. In each of them, waves with transverse and longitudinal wave components are observed.

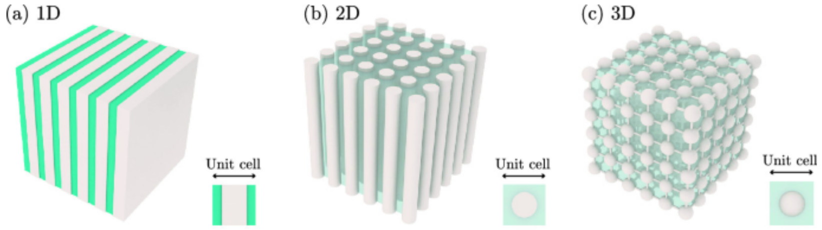
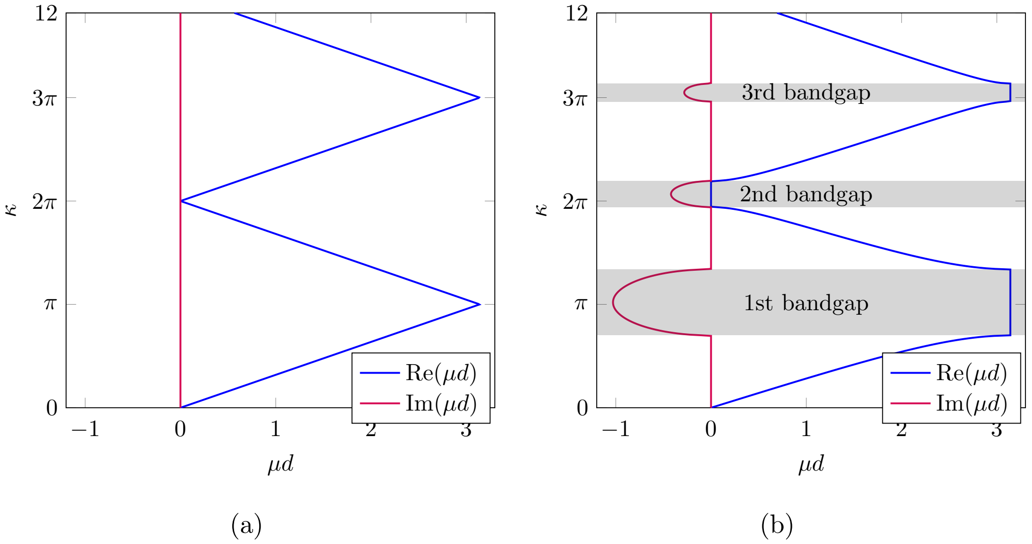


Figure 1.2: Example of simple phononic crystals (from [12]). a) 1D, b) 2D, and c) 3D periodic arrangement.



(a)

(b)

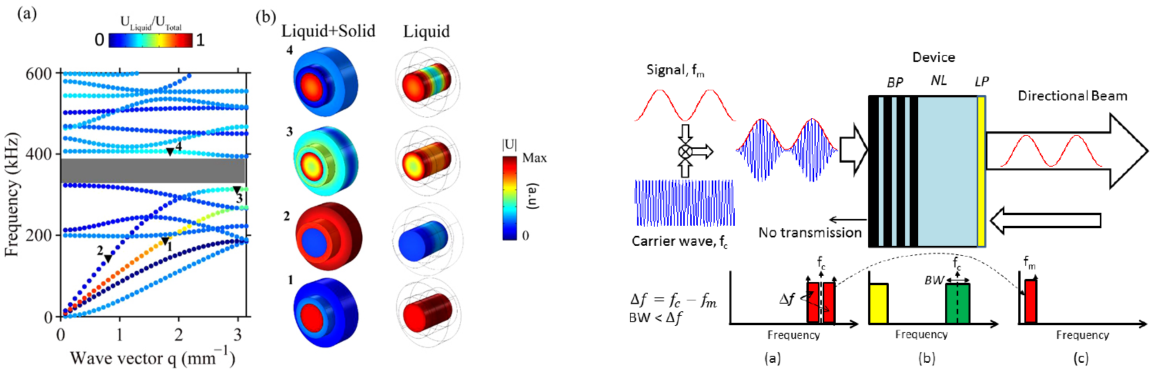
Figure 1.3: a) Dispersion relation without any bandgaps. b) Dispersion relation with bandgaps represented by the shaded regions.

Phononic crystals generally refer to structures, that have the scaterrer made from similar material as the host (e.g. nickel cylinders in copper matrix [7]) and are therefore homogeneous. Whereas the sonic crystals, one of the particular cases of phononic crystals, are structures made of dissimilar materials (e.g. steel rods in fluid). The scatterers are perfectly rigid and therefore sound hard in comparison to the material in which they are placed. Then, only longitudinal wave component can be considered.

The most distinctive representation of the wave transmission through the periodic structures is the dispersion relation, i.e. the relation of the wavenumber to the frequency of the wave. For illustration, Figure 1.3. shows a dispersion relation of a 1D sonic crystal, that will be encoutered later in the thesis. The shaded regions represent the bandgaps. The underlying phenomena behind the formation of bandgaps is Bragg's scattering. Under specific conditions, destructive interference is encountered due to the spatial arrangement of the crystal components, their volume fractions, and the composite materials' sound speed and density ratios [12].

1.3 Applications in medicine and human health protection

Since their beginnings, locally periodic structures have found wide application in many fields of use thanks to their ability to control the propagation of acoustic and elastic waves. In the following we will focus on those applications, that are related to biomedicine and human health protection. The work of Lucklum et al. [13, 14] has shown that phononic crystals can be used for sensing liquid properties in small cavities. The change of material properties inside the phononic crystal leads to changes in the transmission spectrum. Hence, based on the frequencies where the transmission takes place, the properties of the used liquid can be determined.



(a) (b)

Figure 1.4: Examples of sonic crystals applications related to biomedicine and human health protection: a) Tubular phononic crystal (from [15]). b) Superdirectivity filter (from [16]).

The use of phononic crystals for sensing volumetric properties of liquids was further investigated by Gueddida et al. [15]. They designed a tubular phononic crystal (TPC, see Figure 1.4a) where any perturbation of the fluid flow inside the tube is avoided. In this way, the TPC transmission spectrum is sensitive to the mass density and speed of sound variations of the fluid flowing inside the pipe. Their design allows for use in sensing applications, from microfluidics in chemical or biochemical analysis to medical or civil engineering applications.

Applications of phononic crystals were extended by Bonhomme et al. [17] to nanoparticle and bio-molecule sensing. They introduced devices using phononic crystals and Love surface acoustic waves to detect the mass of nanoparticles (e.g. virus, DNA, protein).

The development of phoxonic crystals combining photonic and phononic crystal in one device has opened up the possibility for dual characterization of liquids' acoustic and optical properties. Thanks to its ultrasmall volume and very high sensitivity, it might in future be used as a biosensor, e.g., without the need for immobilization of DNA or proteins [18].

In addition to sensing applications, phononic crystals are also suitable for focalization in areas which rely on high-quality acoustic focusing and imaging, like industrial engineering or medical acoustic imaging. Incorporating phononic crystals or other acoustic metamaterials helps to realize sub-wavelength focusing and super-resolution imaging which is not achievable with conventional lenses [19–21]. Placing of the pnononic crystal in such imaging device can be seen in Figure 1.5.

Moreover, the sonic crystals play a role in superdirectivity needed in medical imaging or underwater communication, where they can provide a band-pass acoustic filter for ultrasound propagation [16] (see Figure 1.4b). Another interesting application is also the thermally tunable phononic crystal lens [22], which allows for adjusting the the focal length by varing the temperature from 20 °C to 39 °C, relying in the temperature-dependent sound velocity and density.

In the context of medical ultrasound, locally periodic structures have another purpose. Where the transducer cannot be coupled to the tissue via a liquid agent, air-coupled ultrasound becomes an alternative. Classical ultrasound may suffer from transmission losses, but recently introduced fluidic transducers (FT) do not need direct contact with the tissue.

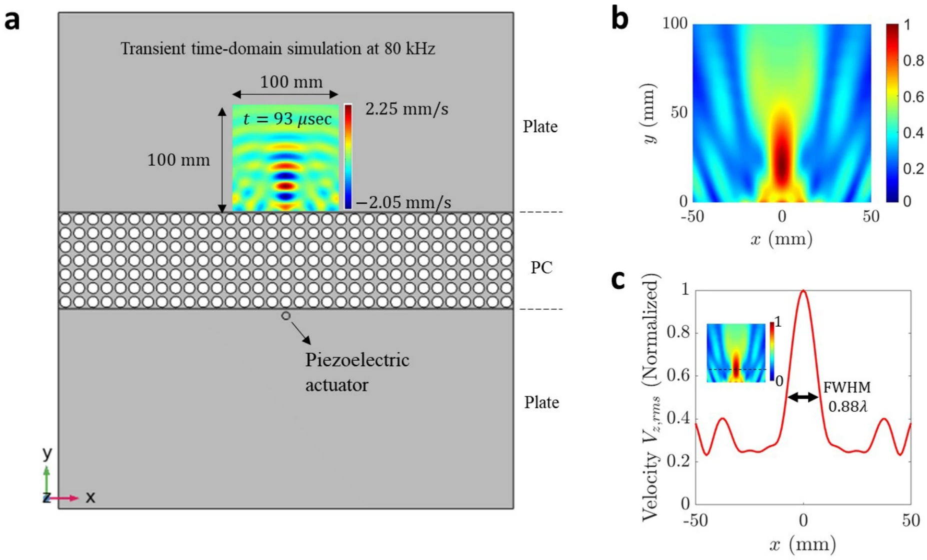


Figure 1.5: Scheme of phononic crystal lens for subwavelength imaging (from [21]).

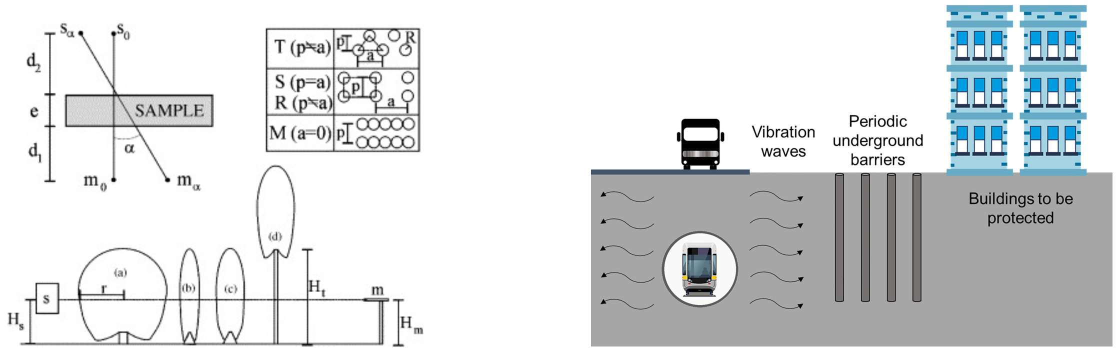
FT uses the instability of supersonic airflow to generate ultrasound. In order to use FT in sensitive applications, the acoustic energy needs to be redirected from the jet axis (the flow field) and for this purpose, sonic crystals could be used [23].

Besides focusing, the application as a frequency filter can be leveraged in sound hygiene, where locally periodic structures can provide an alternative to classic acoustic barriers [12]. The advantage of a sonic crystal acoustic barrier (see Figure 1.6a, Figure 1.6b) is that it is usually lightweight and easy to build. Thota et al. [24] proposed sonic barriers with tunable bandgaps that can be reconfigurable throughout its lifetime and therefore can overcome the issues of fixed periodicity. If the properties of the typical noise source are known, the sonic crystal can be designed in that way, that it attenuates and lets pass what is desired, by alternating the characteristics of the lattice or the scatterers.

(a)

(b)

Figure 1.6: Examples of sonic crystals applications related to biomedicine and human health protection: a) Sonic crystal acoustic barrier (from [12]). b) Periodic pipe noise barrier installed in Eindhoven by Van Campen. (from [24])



(a) (b)

Figure 1.7: Examples of sonic crystals applications related to biomedicine and human health protection: a) Control of noise by trees arranged like sonic crystals (from [25]). b) Isolation of vibration by periodic underground barriers (from [26]).

Especially in transport infrastructures (e.g., railroad corridors, highways), the barriers for traffic noise attenuation or mitigation need to be carefully designed because, in the case of closed classical acoustic barriers, all the emissions stay in the transport corridor and can lead to health and safety hazards, while with sonic crystal acoustic barrier are transparent to air and water [27–29]. The barriers also do not have to be built from fabricated scatterers; the sound attenuation can also be achieved with vegetation: Martínez-Sala et al. demonstrated the effect of trees arranged periodically like sonic crystals [25] (see Figure 1.7a). Similar concept to sonic crystal acoustic barrier is used for attenuation of vibrations generated by traffic or seismic waves: the phononic crystals are buried to mitigate ground wave propagation [26, 30, 31], for illustration see Figure 1.7b.

1.4 State-of-the-art numerical approaches

Approaches used for obtaining acoustic transmission or band gaps in phononic and sonic crystals are the finite difference time domain technique, the transfer matrix method (TMM), multiple scattering theory, plane wave expansion or finite element methods (FEM) [7]. Gupta et al. [32] employ the Webster equation with the Floquet-Bloch theory to study the one-dimensional sonic crystal supporting quasi-plane wave propagation. Recently, the range of methods employed has broadened to incorporate elements of machine learning [33–35].

Although there is a wide variety of numerical methods to solve sonic crystal problems, analytical formulae relating the waveguide geometry and the acoustic transmission are limited to a few special cases, such as the simplest chains (see, e.g., [36, 37]).

When developing devices for sound attenuation or focalization, the design of periodic structures is based on optimization techniques. The wave propagation properties in periodic structures depend on the external shape, the lattice arrangement, the material and the filling fraction [38]. By tuning the mentioned parameters, one can control the acoustic transmission. The optimization process typically involves numerical methods combined with a stochastic search algorithm, such as the method of moving asymptotes [39], gradient-based topology optimization [40], or a genetic algorithm [41]. Nonetheless, this approach is computationally demanding and provides limited insight into the underlying mechanisms.

1.5 Goals of this thesis

Having access to a numerical model is a great opportunity for data-driven discovery and physics-informed machine learning [42–46]. Involving current knowledge of the problem's underlying physics, coordinate transformations, and optimization techniques, formulae relating waveguide geometry and the corresponding dispersion relation can be infered and identified. The aim of this thesis is to propose an analytical model of band structure in one-dimensional sonic crystals of continuous geometries to both improve the overall readability of the system features and enable more efficient optimization. The requirements are that the obtained analytical formulae describing band structure properties of sonic crystals should be physically interpretable and consistent with direct FEM simulations.

The thesis is organized as follows. In chapter 2, the governing partial differential equation is introduced together with the boundary conditions, implications for waveguide geometry and the chosen numerical approach. The proof of concept for discovering analytical model from the data using machine learning is presented in chapter 3. In order to move towards more general case with better interpretability and broader applicability, new dataset is designed and employed with different methods in chapter 4. In chapter 5, the verification of the discovered formulae and illustrative examples of possible application are presented. Finally, the conclusions are drawn in chapter 6.

Chapter 2

Theory

In this chapter the application of governing equations for sound propagation in locally periodic structures is shown. Throughout this work, the case of quasi-planar acoustic waves in a waveguide of a circular transversal cross-section with non-uniform radius is considered. First, the governing partial differential equation is introduced in section 2.1. Its boundary conditions are covered in section 2.2 and the implications for geometry are discussed in section 2.3. Finally, the numerical implementation is described in section 2.4.

2.1 Governing equation

In any acoustic waveguide with variable radius and perfectly rigid walls, the Webster wave equation (see e.g., [47]) holds for the propagation of quasi-plane waves:

$$\frac{\partial^2 p}{\partial s^2} + \frac{1}{A(s)} \frac{\mathrm{d}A(s)}{\mathrm{d}s} \frac{\partial p}{\partial s} = \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} \ , \tag{2.1}$$

where $p = p(s,t)$ is the acoustic pressure and s, t and $A = A(s)$ denote the spatial coordinate along the waveguide axis, time and cross-sectional area function, respectively (see Fig. 2.1). The adiabatic sound speed c_0 is given by $c_0^2 = \gamma p_0/\rho_0$, with γ, p_0 and ρ_0 denoting the ratio of specific heats, ambient pressure and density, respectively.

For simplicity, an axisymmetric waveguide (circular cross-section) is considered and therefore $A(s) = \pi R(s)^2$. Assuming also the time-harmonic behavior (with the sign convention $e^{-\mathrm{j}\omega t}$), the Eq. (2.1) can be rewritten as

$$\frac{\mathrm{d}^2 p}{\mathrm{d}s^2} + \frac{2}{R} \frac{\mathrm{d}R}{\mathrm{d}s} \frac{\mathrm{d}p}{\mathrm{d}s} + \frac{\omega^2}{c_0^2} p = 0 \ . \tag{2.2}$$

In order to express the equation in a non-dimensional form, an axial characteristic length ℓ and a lateral characteristic length R_0 is introduced:

$$\frac{\mathrm{d}^2 \tilde{p}}{\mathrm{d}x^2} + \frac{2}{r} \frac{\mathrm{d}r}{\mathrm{d}x} \frac{\mathrm{d}\tilde{p}}{\mathrm{d}x} + \kappa^2 \tilde{p} = 0 \ , \tag{2.3}$$

where $x = s/\ell$, $r = R/R_0$, $\tilde{p} = p/\rho_0 c_0^2$ and $\kappa = \omega\ell/c_0$ stand for the non-dimensional axial coordinate, local radius, acoustic pressure and wavenumber, respectively.

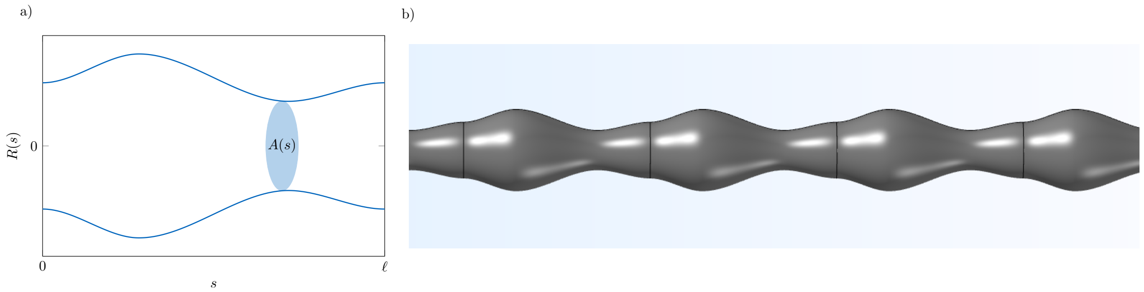


Figure 2.1: a) Waveguide of a circular cross-section with non-uniform radius function $R(s)$ and cross-sectional area function $A(s)$; b) corresponding 3D render from COMSOL Multiphysics.

2.2 Boundary conditions

For the governing partial differential equation with non-constant coefficients, boundary conditions need to be specified. In order to reformulate the problem from wave propagation in an infinite periodic structure to wave propagation in a single unit cell with periodic boundary conditions we assume the radius function $r(x)$ periodic with a period d so that $r(x) = r(x + d)$ and periodic boundary conditions of the Floquet-Bloch type [48, 49]. The Floquet-Bloch theory suggests that

$$\tilde{p}(x + d) = \rho_{1,2}\tilde{p}(x) \ . \tag{2.4}$$

Now the remaining task is finding the multipliers $\rho_{1,2}$. Let $u(x)$ and $v(x)$ denote linearly independent solutions of Eq. (2.3) satisfying the following initial conditions:

$$u(0) = 1 \ , \qquad u'(0) = 0 \ , \qquad v(0) = 0 \ , \qquad v'(0) = 1 \ . \tag{2.5}$$

The characteristic equation for the multipliers then reads

$$\rho^2 - [u(d) + v'(d)]\rho + 1 = 0 \ . \tag{2.6}$$

It follows that $\rho_1\rho_2 = 1$ and hence they can be written as

$$\rho_{1,2} = \exp(\pm \mathrm{j}\mu d) \ , \tag{2.7}$$

where μ stands for the Bloch wavenumber. Moreover, solution of Eq. (2.6) shows that

$$\rho_1 + \rho_2 = u(d) + v'(d) \tag{2.8}$$

and using the substitution in Eq. (2.6) we can write

$$u(d) + v'(d) = \exp(\mathrm{j}\mu d) + \exp(-\mathrm{j}\mu d) = 2\cos\mu d \ . \tag{2.9}$$

Finally, we obtain cosine of the Bloch phase $\xi \equiv \mu d$ and the Bloch wavenumber μ as

$$\cos\xi = \frac{u(d) + v'(d)}{2} \,, \tag{2.10}$$

$$\mu = \frac{1}{d} \arccos \frac{u(d) + v'(d)}{2} \,. \tag{2.11}$$

Although the latter equations are indeed equivalent, we explicitly state both of them for the sake of the sections below.

Generally, the Bloch wavenumber can take complex values ($\mu = \mu(\omega) \in \mathbb{C}$). It follows from Eqs. (2.4) and (2.7) that when the imaginary part of μ is non-zero, the propagation of the quasi-plane wave in the infinite periodic structure is not possible. In a locally periodic structure, only evanescent waves are propagating through.

2.3 Implications for the waveguide geometry

Since the Webster equation is valid in the quasi-plane wave approximation, it is necessary to determine the conditions under which only the quasi-plane waves propagate through the waveguide. The approximation holds when the frequency f of the travelling wave is lower than the cut-off frequency f_c. For an axisymmetric waveguide with the maximum (dimensional) radius $R_{\max}$, the (dimensional) cut-off frequency is given as

$$f_c = \frac{1.83 c_0}{2\pi R_{\max}} \,. \tag{2.12}$$

This cut-off frequency approximation is based on a uniform waveguide (see e.g., [50] for full derivation by perturbation methods) and in practice, the maximum frequency of interest should be significantly below f_c.

The estimate of a characteristic frequency of the first band gap, the Bragg frequency f_B, is given as

$$f_B = \frac{c_0}{2\ell} \,, \tag{2.13}$$

where ℓ is the (dimensional) length of the spatial period. To construct a unit cell which has the first band gap still in the quasi-plane wave region, it has to hold $f_B < f_c$:

$$\frac{c_0}{2\ell} < \frac{1.83 c_0}{2\pi R_{\max}} \,. \tag{2.14}$$

Consequently, the rough estimate of the unit cell aspect ratio must be

$$R_{\max} < 0.58\ell \tag{2.15}$$

which means that the tube diameter must be noticeably smaller than the spatial period.

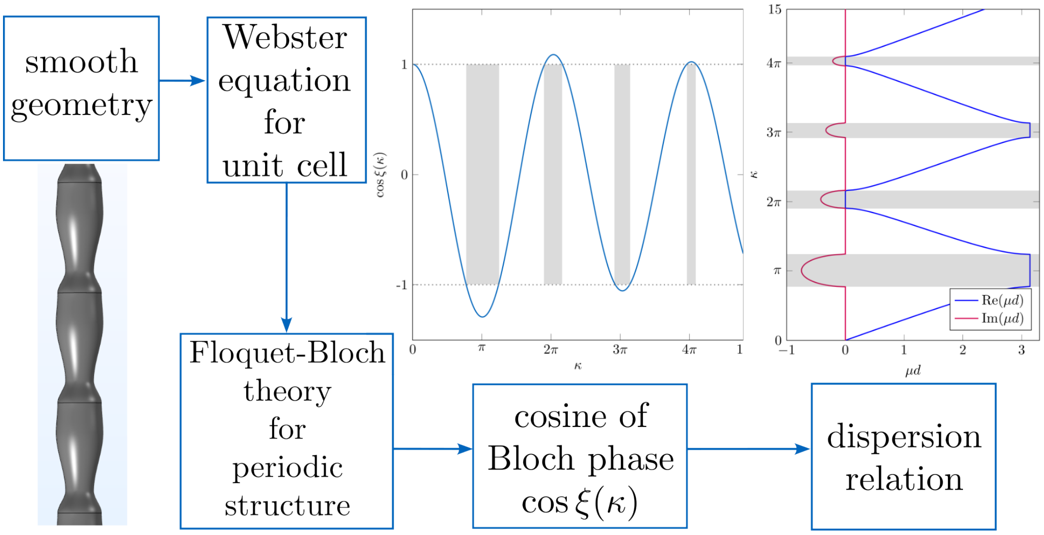


Figure 2.2: From geometry to dispersion relation. The band gaps are depicted in gray.

2.4 Numerical Approach

For a given radius function $r(x)$, the Webster equation (2.3) for a unit cell is solved numerically using Runge-Kutta-Fehlberg method (RK45) with the initial conditions given by Eqs. (2.5). Via the Floquet-Bloch theory, the cosine of Bloch phase $\cos\xi$ is then obtained as explained in section 2.2. To obtain the Bloch wave number μ which is usually displayed in the dispersion relation, arccosine of $\cos\xi$ divided by the non-dimensional period d is taken (see Eqs. (2.10) and (2.11)).

For convenience, the process of obtaining dispersion relation for a given geometry is illustrated in Figure 2.2. The implementation relies on open-source libraries NumPy [51] and SciPy [52].

Chapter 3

Proof of concept

The main purpose of this chapter is to develop a procedure to discover formulae relating waveguide geometry and its corresponding dispersion relation. In this chapter, the employed methods of data-driven discovery will be described. For convenience, a schematic picture of the whole process is shown in Figure 3.1. The dataset design is described in section 3.1. The first step towards the pattern extraction is finding a coordinate system, where the system is simplified (in section 3.2). Section 3.3 deals with the extraction itself (learning equations from data). In section 3.4, the discovered formulae for dispersion relation prediction are revealed, and the results, together with possible limitations of this approach, are discussed in section 3.5.

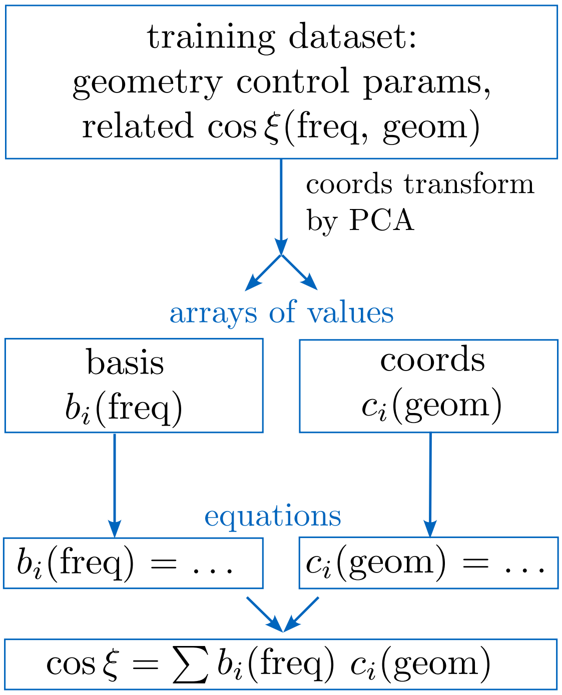


Figure 3.1: Physics-informed machine learning workflow: from the dataset to extracted formulae.

3.1 Dataset design

We aim to discover an analytical model relating waveguide geometry to its dispersion relation. But before we begin to address the problem as formulated, we need to inspect the problem more closely. The Bloch wavenumber $\mu(\kappa)$ plotted in dispersion relation is a non-trivial function taking complex values, which would be challenging to model. Moreover, the transformation of obtaining the Bloch wavenumber $\mu(\kappa)$ from the cosine of Bloch phase $\cos\xi$ is already known: via arccos (see Eqs. (2.10) and (2.11), last step in Figure 2.2). Hence, it is unnecessary to discover it again, and instead of modelling the Bloch wavenumber $\mu(\kappa)$ directly, we simply model the cosine of Bloch phase $\cos\xi$.

3.1.1 Parametrization of the waveguide geometry

Firstly, it is necessary to choose an appropriate waveguide geometry parametrization. The requirements for the parametrization of the radius function $r(x)$ stem from the governing equations of transmission in locally periodic structures. Due to combined prerequisites of the Webster equation validity and the Floquet-Bloch theory, a smooth, slowly varying radius function $r(x)$ is needed, and the unit cells have to be connected periodically and continuously to form a waveguide (i.e., $r(x=0) = r(x=1)$, see chapter 2). So far, there are no analytical solutions to this kind of profiles that can be expressed in terms of basic functions (for nontrivial variants see e.g., [53, 54]). Therefore, there is no reason to prefer any specific type of radius function. On top of the aforementioned requirements, it is essential for further data-driven discovery that the impact of control parameters on the radius function $r(x)$ is clear and easy to interpret.

The specific choice of the unit cell radius function $r(x)$ can be arbitrary, as long as it meets the required conditions. For the beginning, it was decided to choose the following parametrization with two control parameters:

$$r(x) = 1 + A\sin^2[\pi g(x)] \ , \tag{3.1}$$

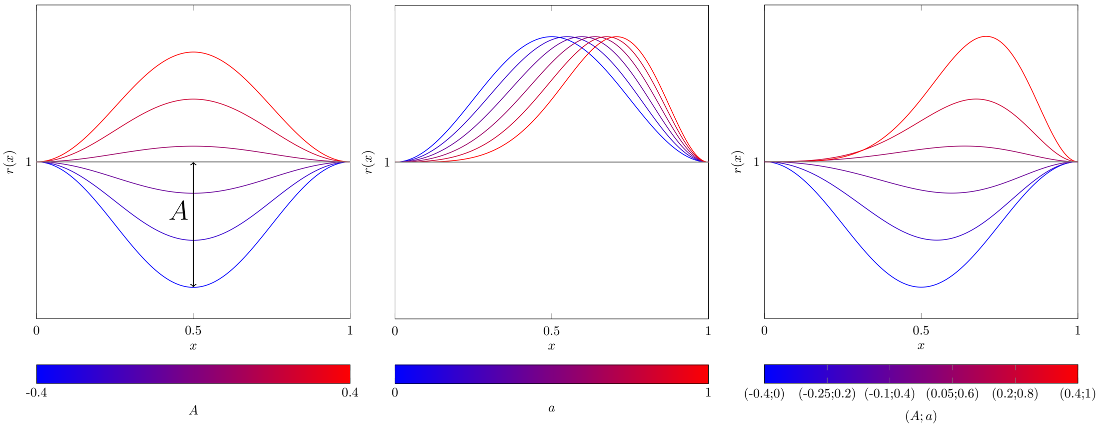
where

$$g(x) = ax^2 + (1-a)x \ , \tag{3.2}$$

and $x \in [0,1], a \in [0,1], A \in [-0.4, 0.4]$ is the spatial coordinate along the waveguide axis, the asymmetry and the amplitude of waveguide geometrical perturbations, respectively. The influence of the two control parameters is shown in Figure 3.2.

3.1.2 Building the dataset

Since the waveguide geometry is uniquely described by the geometry control parameters amplitude A and asymmetry a, the dataset can consist only of these control parameters a, A and related cosines of Bloch phases $\cos\xi(\kappa)$. The range for non-dimensional frequencies κ is chosen as $\kappa \in [0, 15]$ to include the first four bandgaps. The center of the fourth bandgap occurs roughly at $\kappa = 4\pi$, the center of the fifth one at $\kappa = 5\pi$. To be sure that the entire fourth bandgap is present in the computed cosines of Bloch phases $\cos\xi(\kappa)$, the



a) $A \in [-0.4, 0.4], a = 0$ b) $A = 0, a \in [0, 1]$ c) $A \in [-0.4, 0.4], a \in [0, 1]$

Figure 3.2: Influence of geometry control parameters amplitude A and asymmetry a on the waveguide geometry parametrized by Eq. (3.1). For reference, a narrow waveguide $(A = 0, a = 0)$ is depicted in gray.

$\kappa = 5\pi \approx 15.7$ was taken as a number and rounded down. Using the current numerical solution (described in section 2.4), in total 246 cosines of Bloch phases $\cos\xi(\kappa, a, A)$ for given non-dimensional frequencies $\kappa \in [0, 15]$ and waveguide geometries given by a, A was generated (see Figure 3.3).

3.2 Coordinate transformation

To find and extract patterns in data, we first need to find a lower dimensional space, where the system is simplified. This is nothing new in scientific discoveries, where coordinate transformations such as Fourier transform have long been used to reduce complexity of systems. To find our lower dimensional space, we employ a widely used method for data reduction, principal component analysis (PCA).

By applying PCA on the cosines of Bloch phases $\cos\xi(\kappa, A, a)$ from the dataset, we obtain new basis vectors b_i (also called principal components) of a new, lower-dimensional space and new coordinates c_i. The bases b_i are arrays of values shared by the entire dataset, that depend only on the non-dimensional frequency κ $(b_i = b_i(\kappa))$. Meanwhile, each of the coordinates c_i is a scalar distinct for each geometry and hence, the dependency is $c_i = c_i(A, a)$. Then, the original $\cos\xi$ can be expressed as

$$\cos\xi(\kappa, A, a) = b_0(\kappa) + \sum_{i=1}^{n} b_i(\kappa) c_i(A, a) \ , \tag{3.3}$$

where b_0, n is the mean of all cosines of Bloch phases $\cos\xi(\kappa, A, a)$ and the number of chosen bases (as well as the dimension of the new coordinate system), respectively.

For this dataset, the first two dimensions preserve enough variance in the model: 94.62 % and 5.36%, respectively (both can be seen in Figure 3.4). Hence, in this case, we can

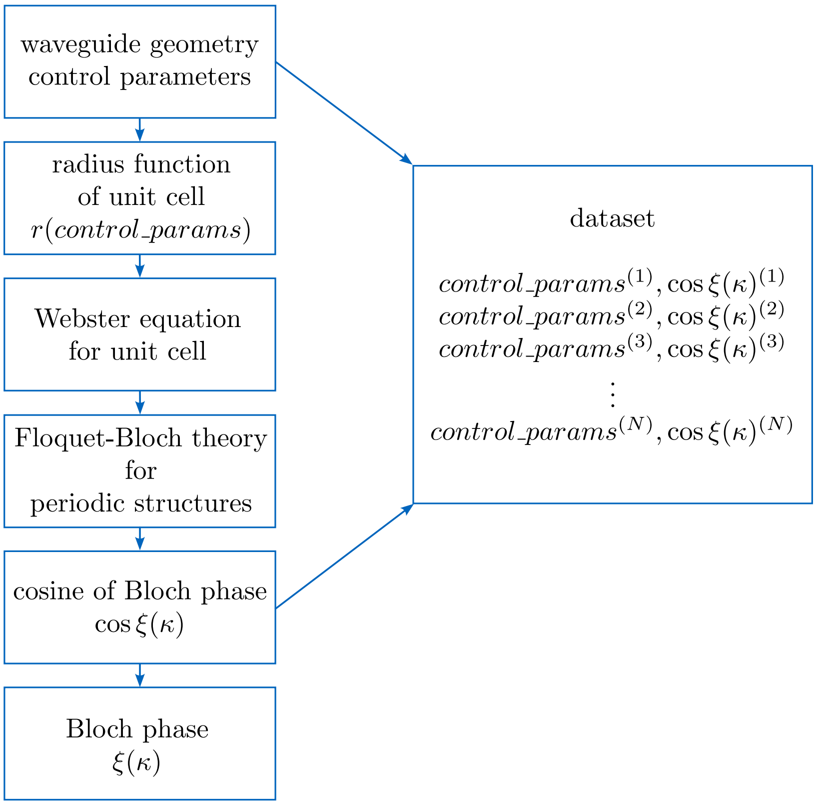


Figure 3.3: Building the dataset.

express the cosine of the Bloch phase as

$$\cos\xi = b_0 + b_1 c_1 + b_2 c_2 \ , \tag{3.4}$$

while keeping 99.98 % of the original information.

3.3 Pattern extraction in the lower dimensional space

The next step towards a model relating geometry to the cosine of Bloch phase $\cos\xi$ is extracting formulae from obtained arrays of values. Each basis function b_i is an array of values shared by the entire dataset; meanwhile, each coordinate c_i is a scalar distinct for each geometry. This task which is done differently for the basis functions b_i and coordinates c_i due to their different nature.

3.3.1 Bases

Extracting equations for the bases b_0, b_1, b_2 shown in Figure 3.4 is a curve-fitting problem that can be solved by, e.g., the method of least squares or eventually its nonlinear and more robust variant, the Levenberg-Marquardt algorithm (LM). By closer inspection of the bases

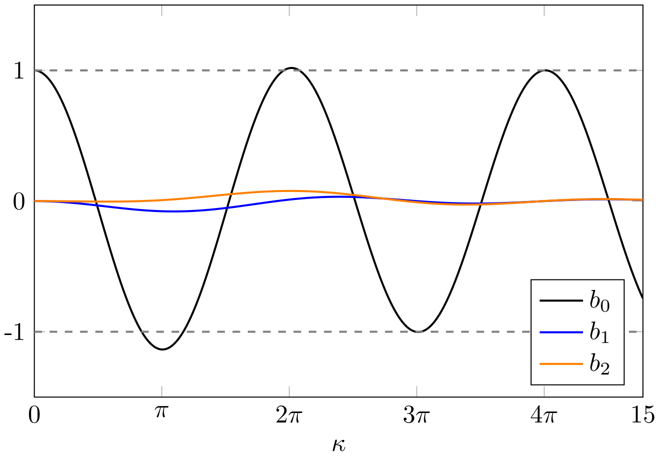


Figure 3.4: PCA. Mean cosine of Bloch phase b_0 and the new basis vectors b_1, b_2, shared by the entire dataset.

b_i in Figure 3.4, we can estimate that the resulting function will probably consist of multiple cosines with different periods and shifted phases, perhaps together with a linear or quadratic polynomial, given as, e.g. Hence, a function can be proposed in the following form:

$$b_i(\kappa) = \beta_1 \cos(\beta_2\kappa + \beta_3) + \beta_4 \cos(\beta_5\kappa + \beta_6) + \beta_6 + \beta_7\kappa + \beta_8\kappa^2 \tag{3.5}$$

where $\beta_i \in \mathbb{R}$ are the fitting coefficients. The reason why LM turned out to be an unsuitable choice is that this algorithm is very sensitive to initial conditions and can be easily stuck in local optima. Another possibility could be using the gradient descent (GD), but that does not solve the task either, since in our case, finding a global optimum with GD is not guaranteed. When using it for our nonlinear curve-fitting problem given by Eq. (3.5) with about eight fitting coefficients, GD also gets easily trapped in local optima without finding the desired fit.

This problem can be overcome by a genetic algorithm, such as Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [55], which is designed for nonlinear non-convex optimization problems. CMA-ES dynamically searches the parameters space, and with a sufficiently large population size (40 candidate solutions in this case), it fits the bases b_i very precisely and quickly.

3.3.2 Coordinates

The relation of coordinates c_1 and c_2 to geometry control parameters amplitude A and asymmetry a is shown in Figure 3.5. are scalars related to each specific radius function $r(x)$. The goal is to relate these scalars as $c_i = c_i(A, a)$. Hence, a library of candidate terms was created, involving asymmetry and amplitude raised to the power of $i = 0, 1, 2, 3$ and their interactions. In order to obtain a sparse, interpretable and stable model, the least absolute shrinkage and selection operator (LASSO)[56] was employed for variable selection.

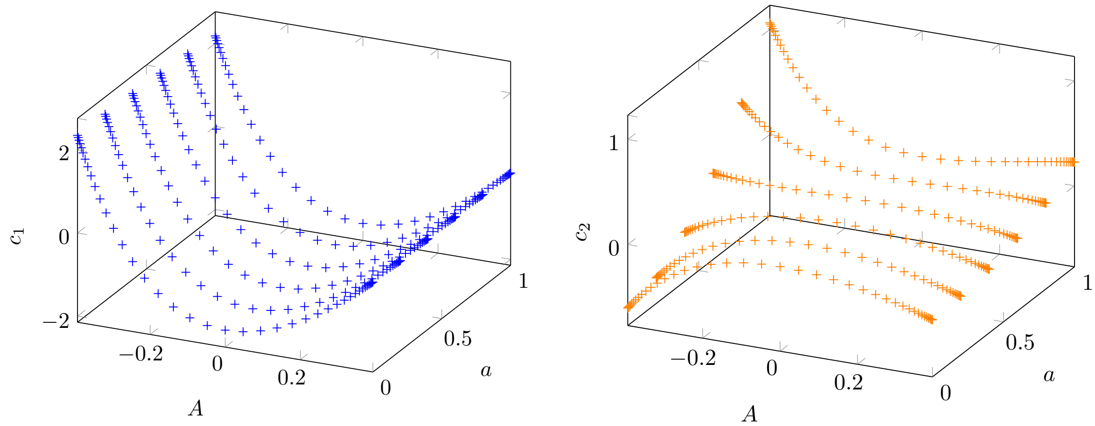


Figure 3.5: PCA. First two coordinates, $c_1(A, a)$ and $c_2(A, a)$, distinct for each geometry.

3.4 Discovered formulae for dispersion relation prediction

The formulae obtained from the described procedure read for the new basis functions

$$b_0(\kappa) = 1.030\cos(\kappa - 0.007) + 0.052\cos(0.675\kappa + 0.659) - 0.073 + 0.007\kappa\ , \tag{3.6}$$
$$b_1(\kappa) = 0.026\cos(0.940\kappa - 0.246) + 0.027\cos(0.634\kappa + 0.933) - 0.040 + 0.003\kappa\ , \tag{3.7}$$
$$b_2(\kappa) = 0.029\cos(0.874\kappa - 0.578) - 0.049\cos(0.362\kappa + 1.377) - 0.017 + 0.003\kappa\ , \tag{3.8}$$

and for coordinates

$$c_1(A, a) = -1.798 + 18.827A^2 - 18.604A^3 + 0.052Aa \tag{3.9}$$
$$+ 0.104Aa^2 - 0.022a^3 + 0.053Aa^3\ , \tag{3.10}$$
$$c_2(A, a) = 0.028 + 0.004A - 2.718A^2 + 2.883A^3 \tag{3.11}$$
$$+ 0.224A^2a - 1.678A^3a + 6.313A^2a^2 - 8.00A^3a^2\ . \tag{3.12}$$

Illustrative comparison of prediction using these formulae with numerical solution for two of the samples from the training dataset can be seen in Figure 3.6. Note that the transmission model is very sensitive to misfitting around $\cos(\xi) = -1$ and $\cos(\xi) = 1$, because of the arccos properties. Although this may not seem like a big issue, it can cause wrong predictions when a band gap should be opened but is not and vice versa. This is visible in example displayed in Figure 3.6 at $\kappa = 3\pi$.

It is necessary to check, if the discovered equations do not lack a physical interpretation. For case of radius function $r(x) = \text{const}$, i.e. $A = 0$ and $a = 0$, we expect $\cos\xi = \cos\kappa$. Therefore, we inspect closely the results for a narrow waveguide depicted in the bottow row of Figure 3.6. Around $\kappa = \pi$ and $\kappa = 2\pi$ the prediction behaves as expected. However, around $\kappa = 3\pi$ the model opens a band gap when it should not, although it is barely noticeable. This limitations of the model together with other are discussed in section 3.5.

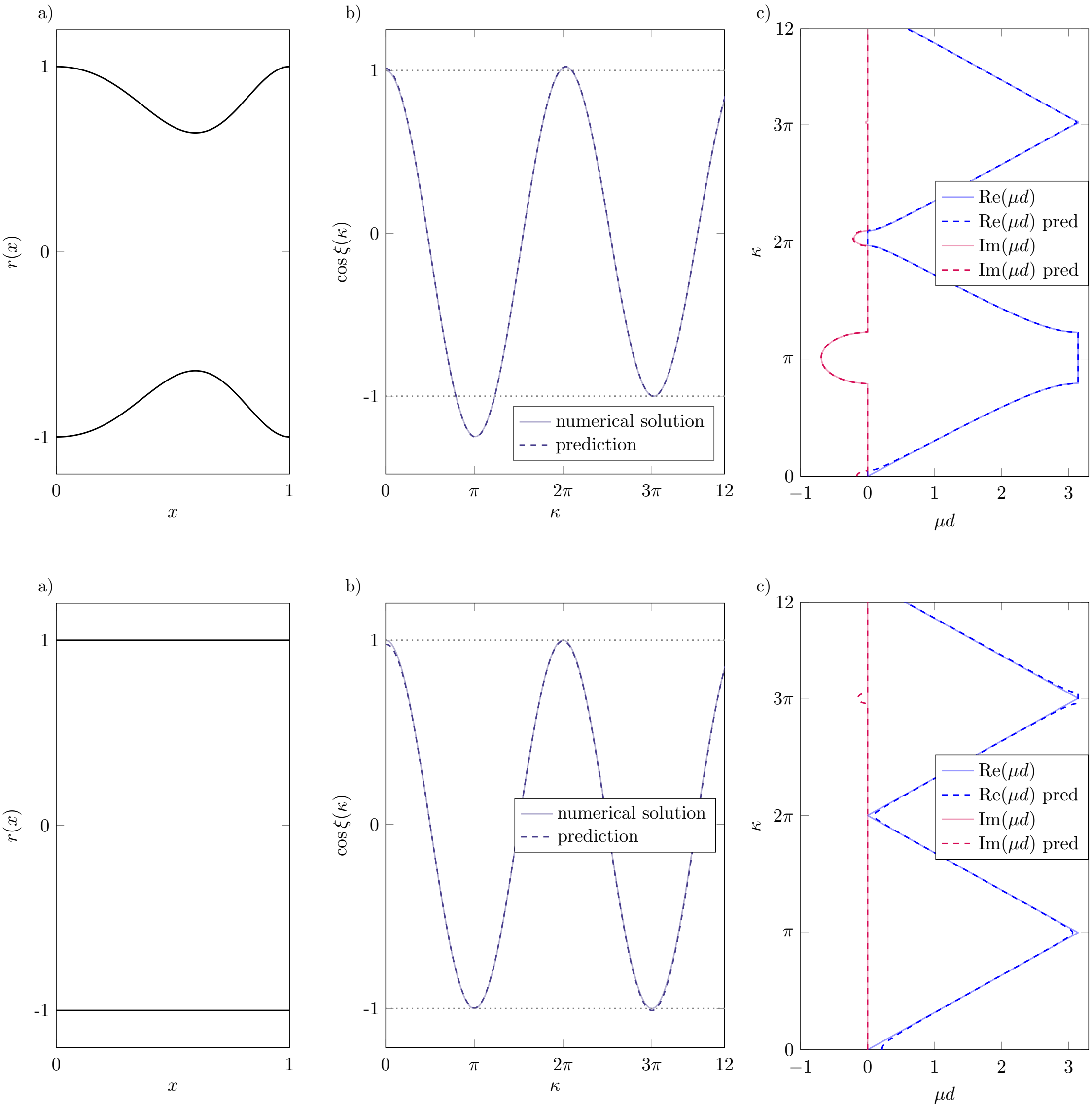


Figure 3.6: Example from the training dataset: a fit (top row), a fit for the narrow waveguide (bottom row). a) Unit cell radius function $r(x)$; b) comparison of the numerical solution and prediction in terms of cosine of Bloch phase $\cos\xi$; c) comparison of the numerical solution and prediction transformed to dispersion relation.

3.5 Discussion

We have verified that it is possible to extract an analytical formula for the Bloch phase from the dataset employing machine learning. To achieve that, a coordinate transformation by PCA was employed. The bases b_i of the new lower dimensional space were fitted employing CMA-ES, and the coordinates c_i were related to geometry control parameters using a candidate library of terms and sparse regression LASSO. Note that the model was not tested on a testing dataset, and no error statistics were made because we encountered two major issues.

Firstly, although the influence of the filling fraction on the band gaps has been studied [12], there is little insight into the relationship between control parameters A, a and the coordinates c_i. When building the library of candidate terms entirely from scratch, it is likely to incorporate some bias based on our intuition and knowledge about the problem. Further extending the candidate library did not turn out to be fruitful; it only led to an exhaustive search. Hence, it is necessary to find a more suitable approach.

Moreover, the chosen geometry parametrization is not variable enough: in most cases, the third band gap is not opened. Hence, a different geometry parametrization with possibly more control parameters is needed, allowing us to create waveguides with more variable radius function $r(x)$ to open and widen the third band gap. Enlarging the dataset and making it more balanced would also help to discover a more accurate relationship between the coordinates c_i and the geometry control parameters because that was the primary source of error so far.

Chapter 4

Towards variability with cubic splines

Throughout this chapter, the proof of concept is further developed to achieve more general and better applicable framework. The new parametrization of a radius function is introduced in section 4.1 and in section 4.2, the used dataset is described. section 4.3 explains and interpretes the results of coordinate transformation applied on this dataset. In section 4.4, it is explained how to extract formulae from the dataset with symbolic regression and in section 4.5, the complete set of formulae is presented. To make design and optimization of sonic crystal even more approachable, section 4.6 focuses on prediction of formulae for bandgap widths and centers. Finally, the results are discussed in section 4.7.

4.1 More general parametrization of the waveguide geometry

As mentioned in subsection 3.1.1, the choice of parametrization is more or less arbitrary as long as it fulfils the requirements that stem from the governing equations of transmission in locally periodic structures: unit cells connected periodically and continuously. Due to the requirements, e.g. higher degree polynomial functions are unsuitable, because they change wildly even for small changes of coefficients and because it is not possible to interprete the coefficients. Since parametrization by Eq. (3.1) with two control parameters does not provide enough variability, it was opted for a cubic spline with four control parameters $x_1, x_2, r_1 = r(x_1), r_2 = r(x_2)$. In addition, it is then easy to require $r(0) = r(1) = 1$, and that the profile has zero slope at the beginning and at the end of the unit cell. For a schematic illustration of the unit cell radius function parametrised by cubic splines, see Fig. 4.1, and for details on the cubic spline, see Appendix in [57].

However, building a dataset of parameters x_1, x_2, r_1, r_2 would lead to discovery of equations, that do not consider that the infinite periodic structures are in this case independent of the mirror symmetry of the unit cell, i.e. the system description is independent of swapping $x_1 \leftrightarrow x_2$ simultaneously with $r_1 \leftrightarrow r_2$.

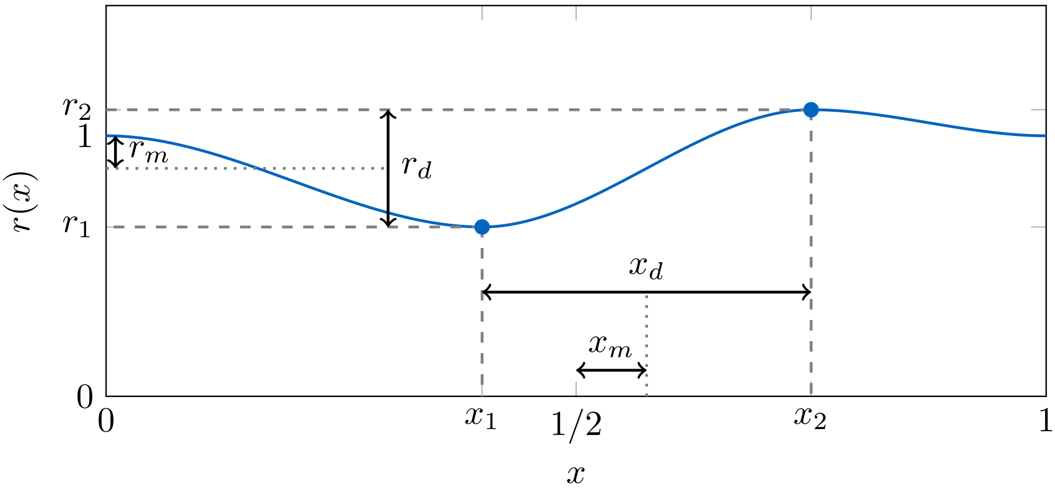


Figure 4.1: Schematic illustration of the unit cell radius function $r(x)$.

Therefore, new control parameters are introduced, that take this symmetry into account:

$$x_m = \frac{1}{2}(x_1 + x_2) - \frac{1}{2} \ , \tag{4.1}$$

$$x_d = |x_1 - x_2| \ , \tag{4.2}$$

$$r_m = \frac{1}{2}(r_1 + r_2) - 1 \ , \tag{4.3}$$

$$r_d = |r_1 - r_2| \ , \tag{4.4}$$

where x_m, x_d, r_m, r_d stand for the excentricity of the average control position, distance between the control points on x, average deviation from the waveguide inlet width, and distance between the control points on r, respectively. For illustration, a narrow waveguide can be described in multiple ways, as long as the average deviation from the waveguide inlet width, i.e. $r_m = 0$.

4.2 Dataset properties

Having the parametrization, a dataset can be constructed, consisting of geometry control parameters x_m, x_d, r_m, r_d and related cosines of Bloch phases $\cos\xi(\kappa)$. Due to requirements of the Webster equation, the range of each control parameter has to be restricted and not all their combinations are allowed. Hence, we limit the non-dimensional slope of the sections connecting the control points and the unit cell endpoints to be lower than 2 (e.g., $|r_1 - r_2|/|x_1 - x_2| < 2$). For the ranges of each control parameter see Table 4.1.

The set of control parameters is created as a meshgrid of all parameter combinations within the given ranges. Subsequently, the combinations are filtered as mentioned above. In this way, a training dataset of 12432 geometries and testing dataset of 4632 geometries were designed together with corresponding cosines of Bloch phases $\cos\xi(\kappa)$ for $\kappa \in [0, 12]$.

Par.	Range	Formula	Interpretation
x_1	$\left[\frac{1}{4}, \frac{9}{20}\right]$	–	x coord. of the 1st control point
x_2	$\left[\frac{11}{20}, \frac{3}{4}\right]$	–	x coord. of the 2nd control point
r_1	$\left[\frac{1}{3}, 2\right]$	$r_1 = r(x_1)$	r coord. of the 1st control point
r_2	$\left[\frac{1}{3}, 2\right]$	$r_2 = r(x_2)$	r coord. of the 2nd control point
x_m	$\left[0, \frac{1}{3}\right]$	$x_m = \frac{1}{2}(x_1 + x_2) - \frac{1}{2}$	excentricity of the avg. contr. point position
x_d	$\left[\frac{1}{5}, \frac{1}{2}\right]$	$x_d = \lvert x_1 - x_2 \rvert$	distance between the control points on x
r_m	$\left[-\frac{1}{3}, \frac{2}{3}\right]$	$r_m = \frac{1}{2}(r_1 + r_2) - 1$	avg. deviation from the waveguide inlet width
r_d	$\left[0, \frac{3}{4}\right]$	$r_d = \lvert r_1 - r_2 \rvert$	distance between the control points on r

Table 4.1: Summary of the control parameters for the unit cell geometry. Note that not all their combinations are allowed due to the requirements of the Webster equation.

Note that in the previous chapter, the range was chosen as $\kappa \in [0, 15]$ to include first four bandgaps. However, during the study, it was realized, that the fourth gap cannot be efficiently controlled by the changes in our parametrization of geometry. This holds also for the new parametrization: more degrees of freedom in the parametrization would be needed to independently control the fourth gap. Hence, the range was chosen like that so it is certain that first three bandgaps are included. The center of the third bandgap is roughly expected to be at $\kappa = 3\pi$, the center of the fourth bandgap at $\kappa = 4\pi$. To be sure that the entire third bandgap is present in the computed cosines of Bloch phases $\cos \xi(\kappa)$, the $\kappa = 4\pi \approx 12.56$ was taken as a number and rounded down.

4.3 Coordinate transformation

To extract formulae for prediction of cosine of Bloch phase $\cos \xi(\kappa, x_m, x_d, r_m, r_d)$ from the dataset, we apply PCA on the dataset to find a coordinate transform into lower-dimensional space that simplifies the system. By doing so, we find new bases vectors $b_i(\kappa)$, dependent only on the dimensionless wavenumber κ, and new coordinates $c_i(x_m, x_d, r_m, r_d)$, dependent only on the radius function features. For this dataset, the first three principal components preserve 99.99 % of the original information (the explained variance) in the model (92.25 %, 7.35 % and 0.39 %, respectively). Hence, we can express the cosine of the Bloch phase in this case as

$$\cos \xi = b_0 + b_1 c_1 + b_2 c_2 + b_3 c_3 \ . \tag{4.5}$$

The bases are depicted in Fig. 4.2. Since it is not possible to show the relation of coordinates to geometry control parameters as in the previous chapter with $c_1(A, a)$ and $c_2(A, a)$ in Figure 3.5 (now we would need to depict it in five dimensions), only their distributions are studied, see histograms in Figure 4.3.

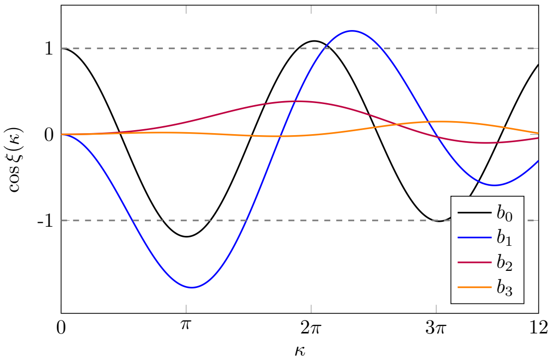


Figure 4.2: PCA results, mean cosine of Bloch phase b_0 and the new bases vectors b_1, b_2, b_3 shared by the entire training dataset.

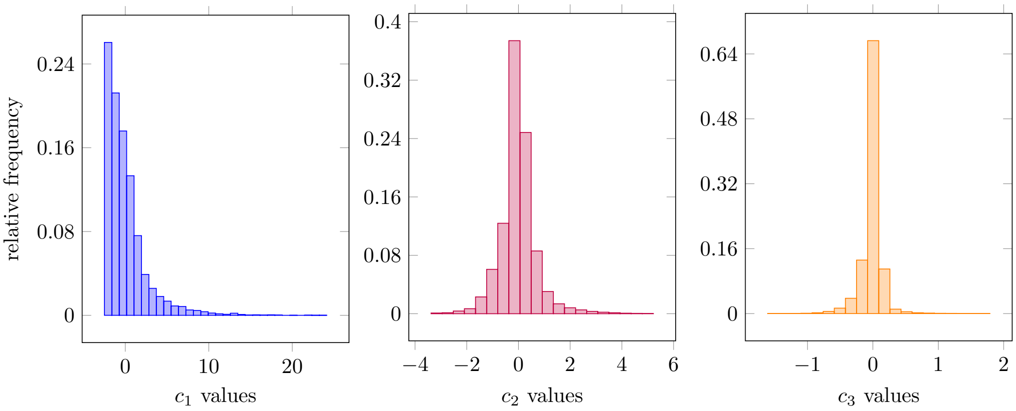


Figure 4.3: PCA results, histograms of coordinates c_1, c_2, c_3.

These histograms were used for some dataset engineering when designing the dataset, because the less outliers the coordinates distributions have, the better the fit could be. The early versions of the training dataset had some wild distributions of both c_2 and c_3, and it was very hard to fit it them to accuracy, that would be satisfactory.

The shapes of bases in Fig. 4.2 are also noteworthy. The mean cosine of Bloch phase b_0 does not correspond to a narrow waveguide (for which holds $\cos\xi = \cos\kappa$), which, in theory, it would be nice. Then all the principal components would describe only the perturbation of the narrow waveguide. Unfortunately, this is not the case. In the beginnings, the training dataset was designed to by symmetrical in geometry (because that was the only symmetry possible to control directly) but the underlying physics is not symmetrical (as will be discussed later in section 4.6). Hence, it did not provide any advantage and the efforts were more fruitful when focusing on the coordinates distributions while designing the dataset.

The base b_1 is essentially the only one active in the region around $\kappa \approx \pi$, i.e., in the region of the first bandgap. It is also responsible for the shift of the second bandgap towards higher κ. The second base b_2 is responsible for widening the second bandgap and the third base b_3 is mainly active in the region of the third bandgap.

4.4 Towards suitable form of regression

Following the approach introduced previously in section 3.3, PCA is applied on the cosines of Bloch phases $\cos\xi(\kappa)$ and we obtain basis functions b_i and coordinates c_i, which need to be related to the non-dimensional wavenumber and geometry control parameters, respectively.

If one would like to follow the approach even further, it would include building own library of candidate solutions to find formulae in the form of $c_i = c_i(x_m, x_d, r_m, r_d)$. Previously, the candidate terms were the geometry control parameters raised to the power of $i = 0, 1, 2, 3$, allong with all possible interactions between them. For two control parameters, this results in a total of 24 candidate terms. For four control parameters, the library includes 112 terms and it takes the LASSO significantly more time to converge. In the end, applying LASSO on this library candidate solutions did not turn out to be successful.

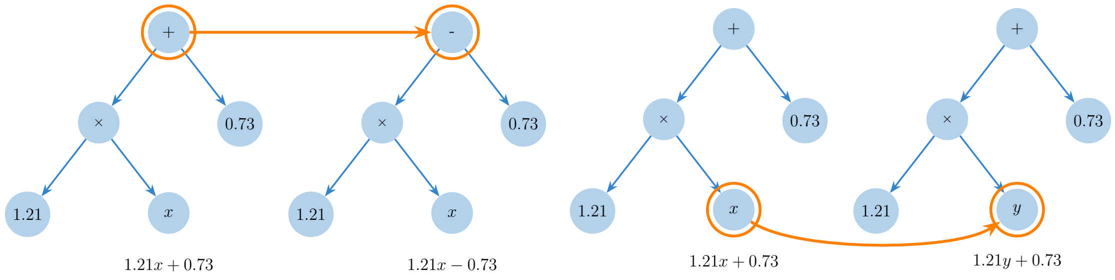
This data-driven discovery would benefit from including many more different candidate terms, but that would even enlarge the library and still would not guarantee any success. One key issue lies in the bias introduced by hand-selecting candidate terms, which inherently limits the model space. In essence, with only minimal guidance during learning, the core task is to identify expressions that accurately fit the observed coordinates c_i, but are also sparse and interpretable. And this is the central challenge of symbolic regression (SR), which searches the space of mathematical expressions for a model that accurately represents the relationship between the predictors and the dependent variable [58].

One of the recent algorithms, the so called AI Feynman proposed by Udrescu and Tegmark [45], employs neural networks for data interpolation. The trained neural network is then used for recursive breaking the problem into smaller parts by finding simplifying properties, such as hidden symmetries and invariances in the data. E.g. to find the translational symmetry, it is checked whether the input $(x_1, x_2, x_3, ...)$ gives the same result as $(x_1 + a, x_2 + a, x_3, ...)$, for various constants a. If it holds within a given precision that $f(x_1, x_2, x_3, ...) = f(x_1 + a, x_2 + a, x_3, ...)$, it can be concluded, that the function f (yet to be found) depends only on the difference of x_1 and x_2. This means, that the variables x_1 and x_2 can be replaced by a a single variable $x'_1 \equiv x_2 - x_1$. Although this idea seems promissing, applying this algorithm on finding relationship between coordinates and geometry control parameters did not turned out into success.

Finding a symbolic expression that fits data generated by an unknown function has long been considered a computationally intensive and challenging task, often assumed to be NP-hard. This assumption was formally proven in [59], which further justifies the reliance on heuristic methods for SR. One of the most prominent heuristic approaches is genetic programming, popularized for SR by [60], which draws on principles of natural selection.

In this method, solution candidates are represented as expression trees composed of analytical functions and operators. The use of genetic programming, a method based on principles of natural selection, was for SR was popularized by [60]. Solution candidates are built as expression trees from analytical terms and operators. A population of solution candidates is then iteratively evolved in generations by repeated application of selection, crossover, mutation and replacement. Examples of mutations are shown in Figure 4.4, an example of crossover operation in Figure 4.5. This approach presents an advantage: there is no need to make assumptions about the underlying the physics and bias the model, because the model structure is not specified beforehand (as it would be for e.g. polynomial regression) and is learnt on the fly.

Overview of available implementations and libraries is provided by many papers on this topic, see e.g. [61, 62]. Here, an open-source library PySR [63] is employed, which uses Julia for efficiency and Python frontend for user convenience. A key aspect of the library setup is the choice of suitable candidate functions from which the expressions are constructed, together with the size and amount of populations.



(a)

(b)

Figure 4.4: A mutation operation applied to a) an expression tree, b) an input variable.

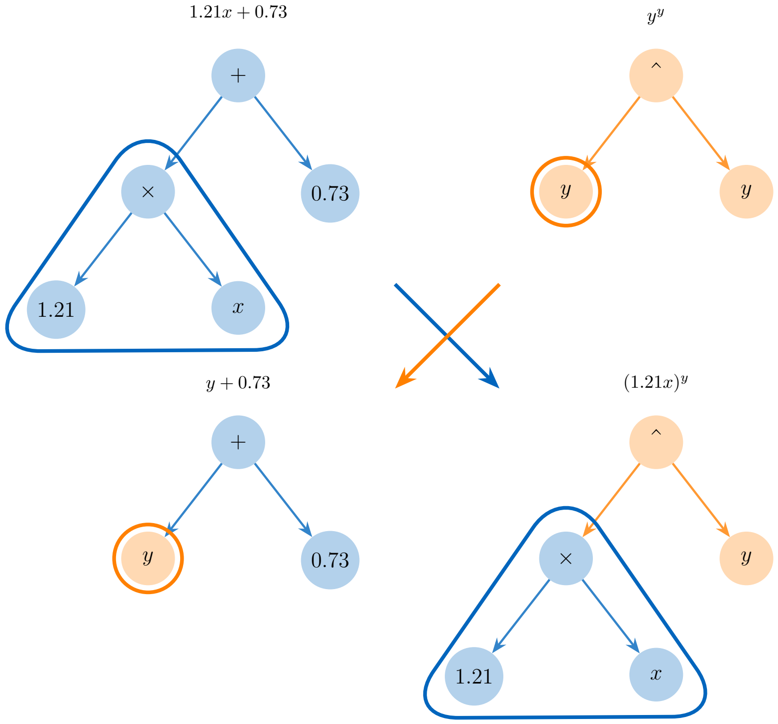


Figure 4.5: A crossover operation between two expression trees.

4.5 Discovered formulae for dispersion relation prediction

After a good deal of numerical experiments, the best choice in terms of simplicity, interpretability and accuracy was to allow the functions sin, cos, exp and the basic binary operators $+, -, *$ and $/$. For fitting the bases b_i (i.e. curve fitting), the default option for loss function was used: the mean squared error `L2DistLoss()`. However, applying the same loss function also for the coordinates fitting did not turned out to be succesful, because of how the points are distributed in space - the mean squared error is sensitive to outliers. Hence, the quantile loss function `QuantileLoss(τ)` was employed instead to mitigate the effect of outliers, which improved the SR performance.

The losses are implemented in the library PySR via package LossFunctions.jl, where the quantile loss function is defined as

$$L_\tau(r) = \begin{cases} (1-\tau)r & \text{if } r \geq 0 \\ -\tau r & \text{otherwise} \end{cases} \tag{4.6}$$

with $r = \hat{y} - y$. The goal was to focus on the central tendency, so the parameter τ was chosen accordingly, to be $\tau = 0.5$.

The formulae for bases functions $b_i(\kappa)$ were fitted with the use of PySR as:

$$b_0(\kappa) = 1.054\cos(\kappa) - 0.115\sin(0.587\kappa) - 0.026 \ , \tag{4.7}$$

$$b_1(\kappa) = 1.134\cos(0.883\kappa) + 0.193\kappa - 1.279 - 5.8 \cdot 10^{-6}\exp(\kappa) \ , \tag{4.8}$$

$$b_2(\kappa) = -0.037\kappa\cos(0.617\kappa) + 0.032\kappa - 1.6 \cdot 10^{-6}\exp(\kappa) \ , \tag{4.9}$$

$$b_3(\kappa) = 0.009 \ \kappa \ [\sin(0.857\kappa) - \sin(-0.105\kappa)] \ . \tag{4.10}$$

The quality of the predictions can be seen in Figure 4.6. While for b_0 and b_2 the prediction is almost perfectly follows the outputs of PCA, some discrepancies occur for b_1 and b_3, which was a tradeoff between accuracy and sparsity of the formulae.

For the coordinates $c_i(x_m, x_d, r_m, r_d)$, the fitted equations read:

$$c_1 = -0.134\cos(2.990r_m) + \frac{r_d^2 + 0.073}{3.894r_m + 2.448 - r_d} \ , \tag{4.11}$$

$$c_2 = -(9.196x_m + 0.305)\frac{r_m r_d}{r_m + 0.529} \ , \tag{4.12}$$

$$c_3 = 0.033 - 0.710x_m\left(\frac{x_m}{x_d^2} + 2r_d\right)\left(r_m + \frac{r_d}{r_m + 0.390}\right) \ . \tag{4.13}$$

Note the formula for the coordinate c_1 (Eq. (4.11)) which is independent of x_m and x_d, and depends only on the relative addition or removal of the waveguide breadth. Considering that c_1 is dominant mainly in the region of the first bandgap, we can conclude, that the first bandgap is influented by the change of unit cell overall volume, not its shape (for the reasoning, see the end of Appendix in [57]).

In Figure 4.7, the prediction versus truth plots for Eqs. (4.11) - (4.13) are shown. In the left column, the overall trends of coordinates and their predictions are shown. As discussed before, it is not feasible to plot the dependencies of coordinates on the gemetry control

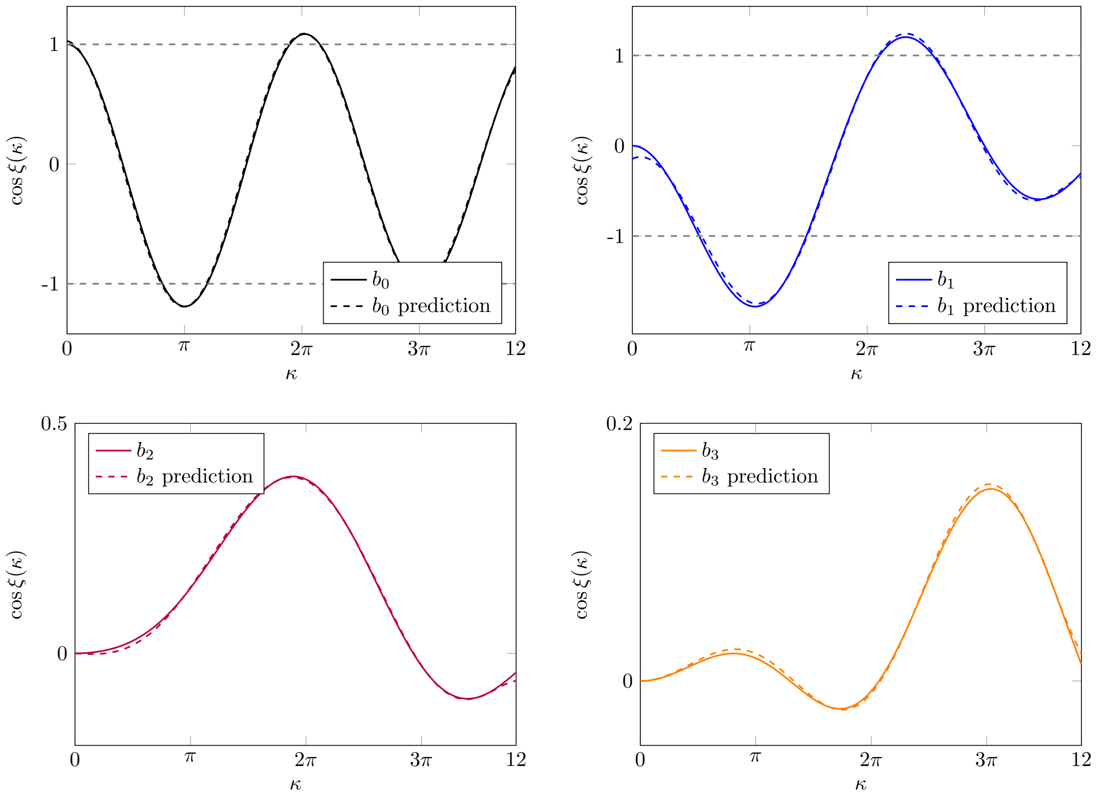


Figure 4.6: Bases and their prediction, based on Eqs. (4.7) - (4.10).

parameters. Hence, they are plotted in the order how they show up in the training dataset, which is sufficient for rough orientation.

Based on the prediction versus truth, the model for coordinate c_1 performs well: the points are either very close to or directly lie on the line given by prediction equal to truth. However, this is not the case for the two other coordinates c_2 and c_3: they show higher prediction error and variance. The higher error variance could indicate systematic underfitting. Although allowing more complex expressions could enhance the performance, they would no more be as sparse and interpretable. The choice of the presented set of equation reflects the tradeoff between accuracy, sparsity, and interpretability.

The bases and coordinates are put together and the formula for the cosine of Bloch phase then follows as

$$\cos\xi(\kappa, x_m, x_d, r_m, r_d) = b_0(\kappa) + \sum_{i=1}^{n=3} b_i(\kappa) c_i(x_m, x_d, r_m, r_d) \ . \tag{4.14}$$

We are interested in comparing the cosine of Bloch phase $\cos\xi$, where the ground truth is given by the numerical solution and the prediction by the discovered formulae. Although calculating either the mean absolute error or the mean squared error is certainly a way, we are far more interested in the maximum difference between the cosines of Bloch phase $\cos\xi$. For the comparison of this error on the training and the testing dataset see Figure 4.8. As

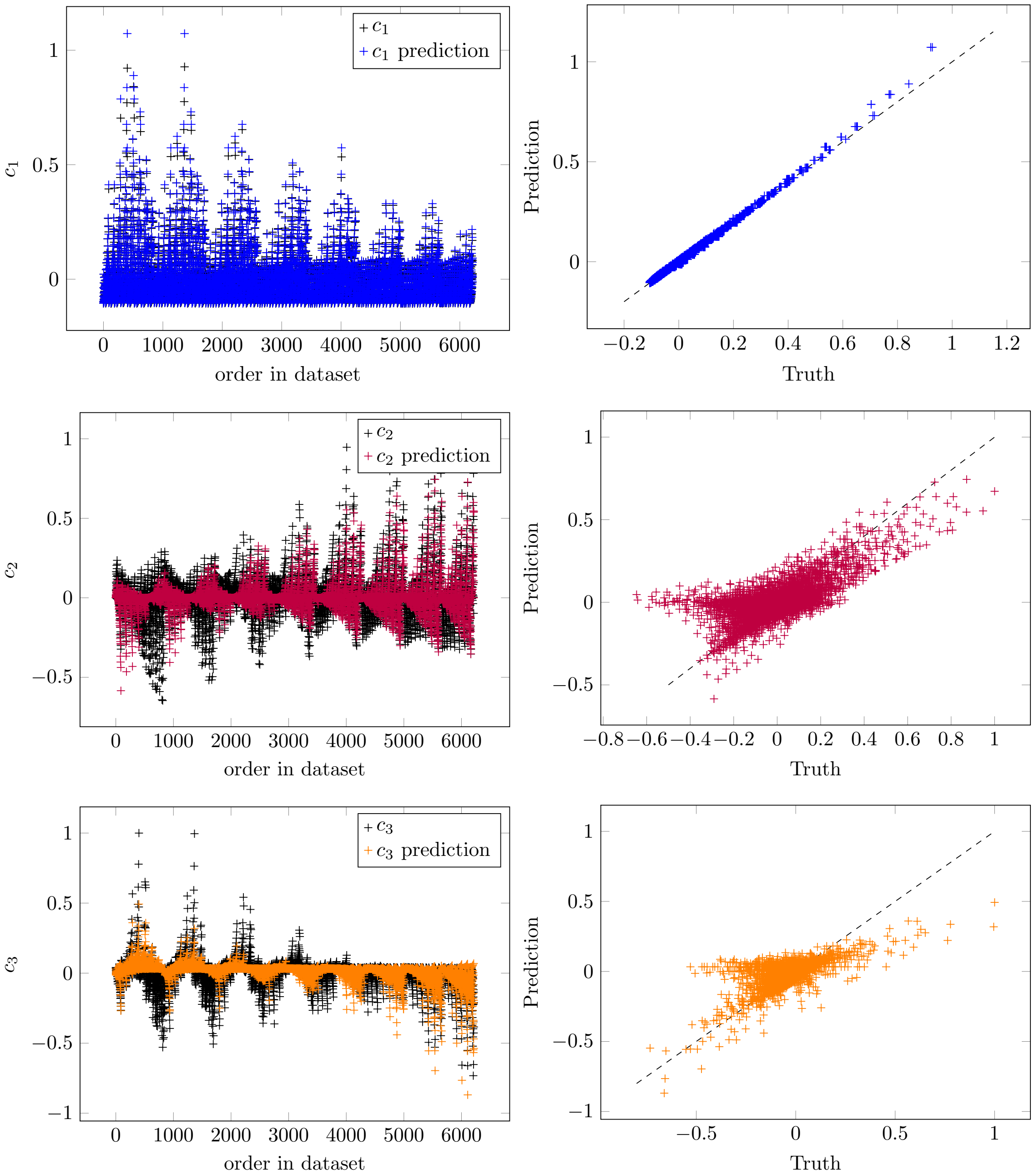


Figure 4.7: Coordinates and their predictions, based on Eqs. (4.11) - (4.13).

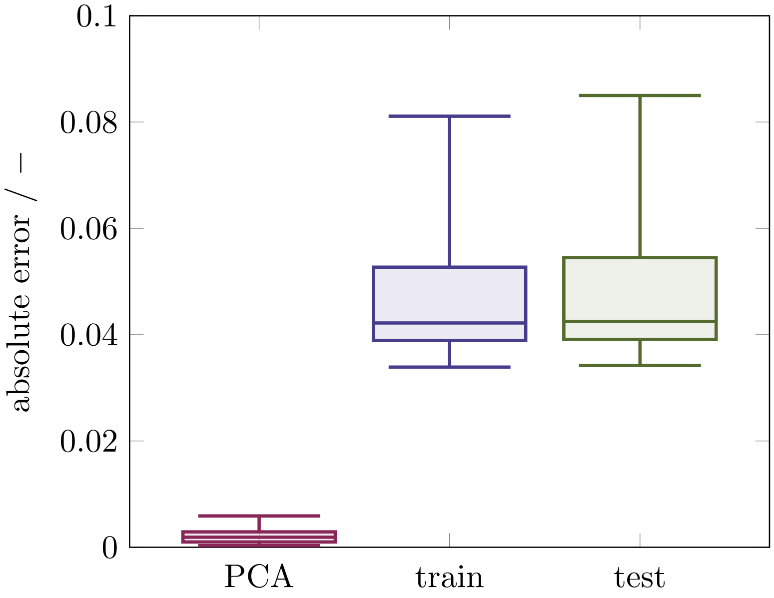


Figure 4.8: Error on the training and the testing dataset. The error was obtained as the maximum difference between numerical solution and prediction in terms of cosine of Bloch phase $\cos\xi$.

mentioned before, the fit of both b_i and c_i is a trade-off between accuracy of prediction and simplicity of the extracted formulae, while still some of the formulae might be underfitted. Significantly reducing the error on the training dataset is not possible in case we want to keep the interpretability of the model. On the other hand, the error on the training and the testing dataset are comparable, therefore we can conclude, that no overfitting is present.

Although negligible, the first error has been introduced into the system already at the beginning with the coordinate transform: as described in section 4.3, by choosing to use the first three principal components, 99.99 % of the original information is preserved (and 0.01 % lost).

A few examples from the training dataset are shown in Figure 4.9, where the numerical solutions of the Webster equation used for training are compared to the extracted formulae. The first case illustrates that not every prediction fits perfectly as the second case with no difference between the numerical solution and the prediction. The last case shows a narrow waveguide, where it looks like $\cos\xi = \cos\kappa$ mostly holds and one can begin to celebrate that the discovered formulae do hold for the basic case without any perturbances in geometry. Unfortunately, this case also demonstrates that the transmission model is very sensitive to misfitting around $\cos\xi = -1$ and $\cos\xi = 1$: although not large, there are visible differences between numerical solution and prediction in the dispersion relation.

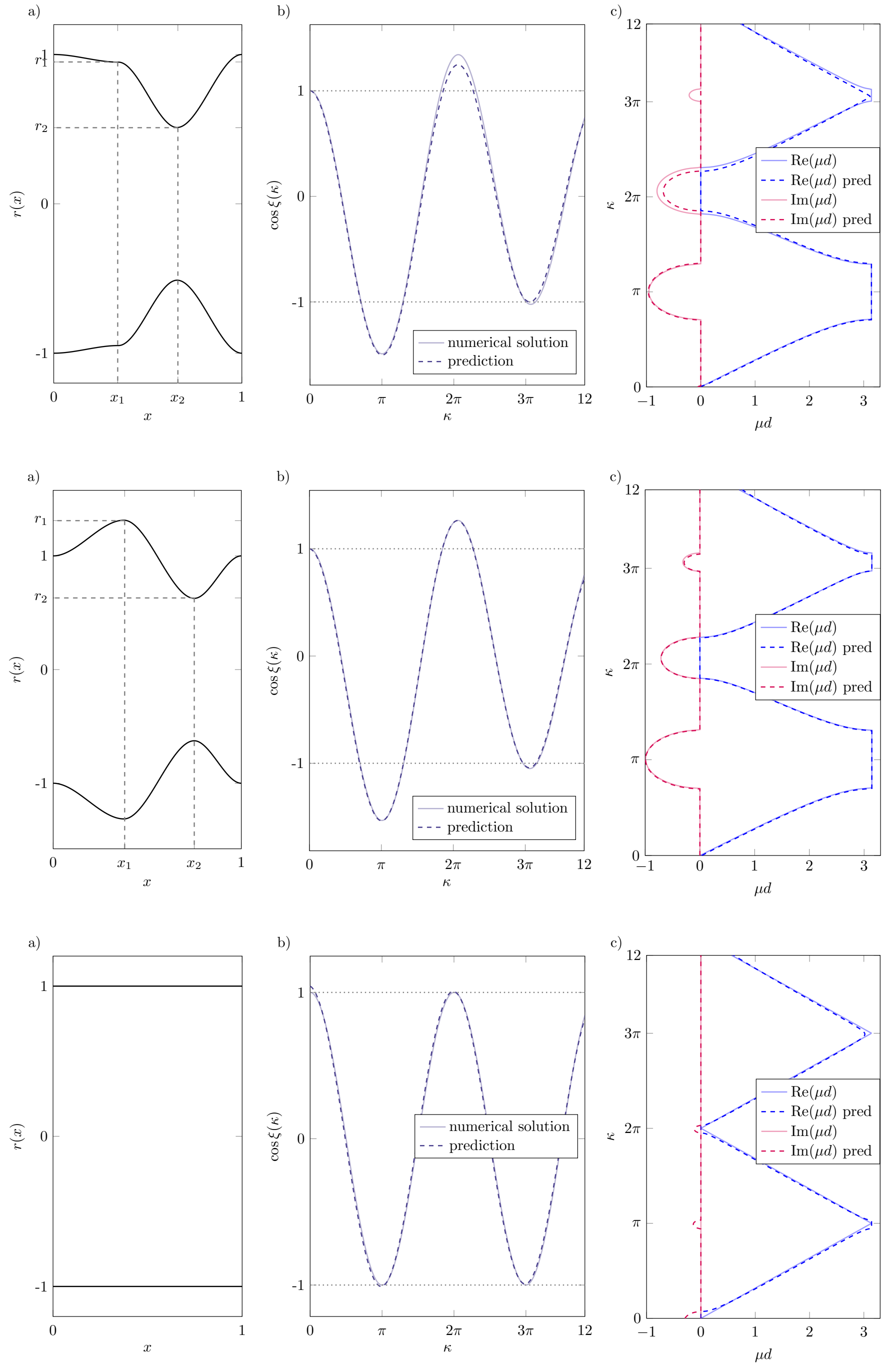


Figure 4.9: Examples from the training dataset: a fit (top row), a good fit (middle row), a fit for the narrow waveguide (bottom row). a) Unit cell radius function $r(x)$; b) comparison of the numerical solution and prediction in terms of cosine of Bloch phase $\cos\xi$; c) comparison of the numerical solution and prediction transformed to dispersion relation.

4.6 Discovered formulae for width and centre of bandgaps

Having the final formula for the dispersion relation prediction is a first step towards having some insight in the system and how it is influenced by the geometry. However, for proposing an optimized design for a desired band gap width, it could be usefull to have formulae describing only that and not the whole cosine of Bloch phase $\cos\xi$.

The first idea on how to predict the bandgap widths was performing Taylor expansion about points given by the first, second and third Bragg frequency (i.e., $\kappa = \pi, 2\pi, 3\pi$, recall section 2.3), which approximately estimate the bandgap centres. Although this might be straightforward in theory, the expressions become too complicated. Hence, this approach would go entirely against the chosen philosophy: improving the readability of the system by discovering sparse and interpretable formulae. The SR has proven suitable for tackling our problem of fitting bases b_i and coordinates c_i, and hence, it was decided to employ it to find formulae for the width and centre of bandgaps.

The centres m_1, m_2, m_3 and the widths w_1, w_2, w_3 were retrieved from the datasets in the following manner. The width of bandgaps are extracted from the $\cos\xi$ as the width of range where $\cos\xi \leq -1$ for the first and the third gap and $\cos\xi \geq 1$ for the second gap. The centers lie in the middle of that range.

After a good deal of numerical experiments, the best choice in terms of simplicity, interpretability and accuracy was to assemble the formulae for centers and widths from functions min, max and binary operators $+, -, *$ and $/$. To reduce the importance of outliers, the L1-loss function was employed (i.e., the mean absolute error).

Since the Bragg frequencies give a rough estimation of the centres of bandgaps, it makes sense to describe only the deviations from these values. The fits provided by PySR read

$$m_1 = \pi + 0.38\max(0.17, x_d r_d) , \tag{4.15}$$

$$m_2 = 2\pi + 0.3\left[\max(r_m, r_d) - \min(0, r_d)\right] , \tag{4.16}$$

$$m_3 = 3\pi + \max(0.25, r_d - r_m)\max(0.3, r_d) . \tag{4.17}$$

The summary statistics of errors for these formulae in Figure 4.10 shows, that the Eqs. (4.15)–(4.17) describe the center frequencies very accurately. However, this is partially because the centers being described do not change much and the Bragg frequency as an estimate is reasonably accurate on its own. Note that this is very likely a feature of the smooth geometries in focus, not of locally periodic structures in general.

On the other hand, the accuracy of formulae for the gap widths found via SR varied greatly, and it was therefore necessary to trade simplicity and interpretability of the formula for its accuracy. Hence, two sets of formulae are shown. First, the ones with lower complexity (denoted with a hat):

$$\hat{w}_1 = 2.61\ \left[\max(r_m, r_d) - \min(0, r_m)\right] , \tag{4.18}$$

$$\hat{w}_2 = 0.13 + \max[1.88(r_d - r_m), 4.67 x_d r_m] , \tag{4.19}$$

$$\hat{w}_3 = 0.53\ r_d - 0.84\min(0, r_m) . \tag{4.20}$$

From these, several noteworthy functional dependencies can be observed. Firstly, the width of the first gap in Eq. (4.18) does not depend on the x-position of the control points, only on

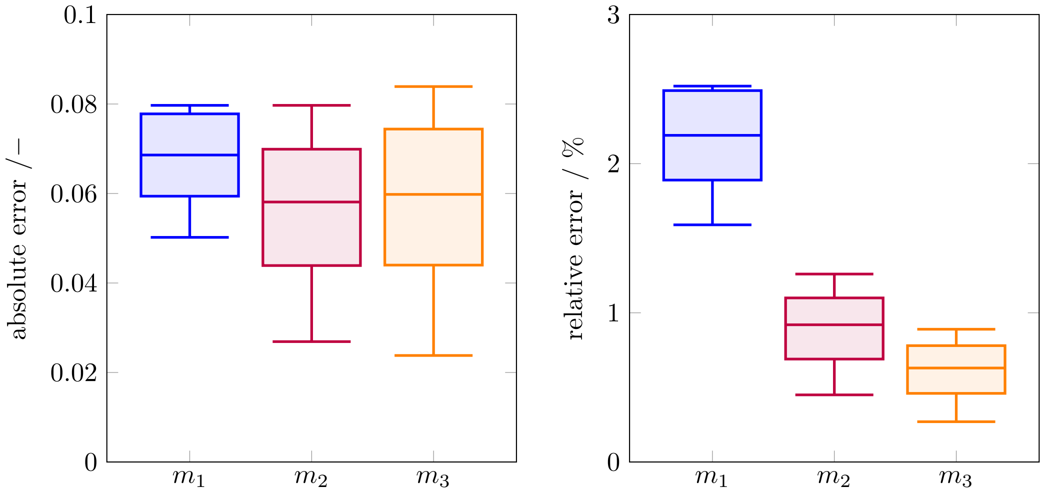


Figure 4.10: Error on testing dataset for bandgap center prediction (Eqs. (4.15)–(4.17)). The box is indicating the median and the interquartile range, the whiskers 8th and 92th percentile of error.

the relative addition or removal of the waveguide breadth. This is consistent with previous findings: the formula for coordinate c_1 (Eq. (4.11)), that is dominant in the region around $\kappa \approx \pi$, (i.e., the first bandgap) contains a dependence on the same parameters as the equation for the width of the first gap (Eq. (4.18)). This confirms the previous implications that the decisive factor for the first gap is the change of unit cell overall volume. Last but not least: the Eq. (4.19) suffers from physical inconsistency: according to that formula, even a straight tube would have a small second gap, while it should have none. One of the future goals would be designing a custom loss function, that would penalize such formulae so we would not have to select the physicaly viable ones by hand.

The comparison of formulae with lower complexity and formulae with higher complexity is shown in Figure 4.11. It is clear that although the lower complexity formulae for w_2 and w_3 might be sufficient for providing general insight into the system, the accuracy is not sufficient for design and optimization. The more elaborated set of equation reads

$$w_1 = 2.61\ [\max(r_m, r_d) - \min(0, r_m)]\ , \tag{4.21}$$

$$w_2 = \max(4.95 x_d r_m,\ r_d) - (r_d + 0.11)\big[(1.85 x_d - 1.01) + 5.62 \min(r_m - x_m,\ x_m)\big]\ , \tag{4.22}$$

$$w_3 = \max[\min(x_d, r_m),\ (r_d - r_m)] \frac{\max(x_d,\ 0.7 - x_d)}{0.37 - x_m} - 0.46\ . \tag{4.23}$$

These equations might be harder to read but still very simple to be implemented in an auxiliary script for evaluation and optimization, moreover useful special cases can be derived from them. Note that Eqs. (4.18) and (4.21) are identical because the simpler relationship was applicable to the same degree of accuracy as the more complex ones.

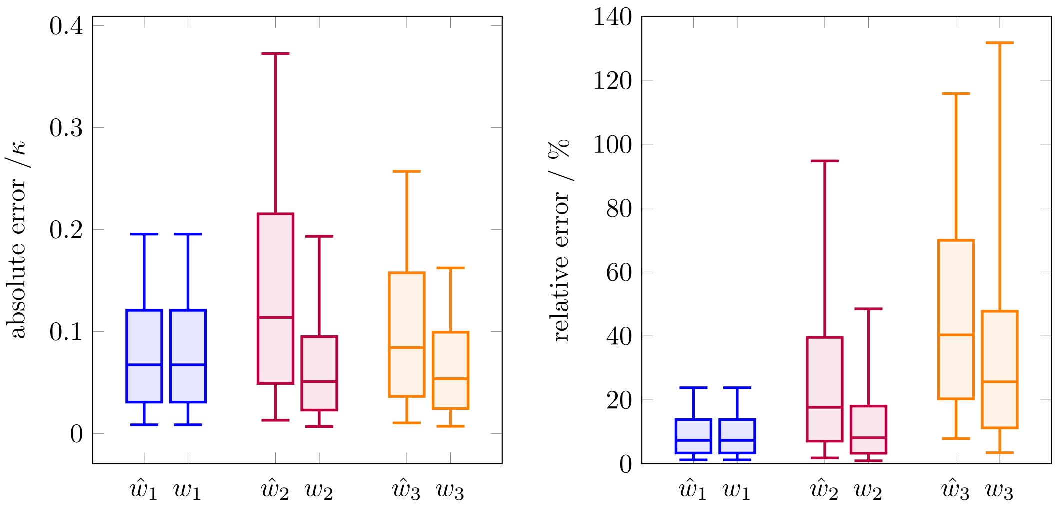


Figure 4.11: Error on testing dataset for bandgap width prediction (Eqs. (4.21)–(4.23)), measured on the κ-axis. The box is indicating the median and the interquartile range, the whiskers 8th and 92th percentile of error. Note that the hat is indicating the results of relations with lower complexity (Eqs. (4.18)–(4.20))

4.7 Discussion

Throughout this chapter, all formulae discovered from the training dataset were compared against the test dataset to check their performance and to rule out the possibility of overfitting. The accuracy remained comparable for both datasets (check, e.g. Figure 4.8) and therefore, it was not necessary to address the overfitting phenomenon.

In comparison to the previous chapter 3, the waveguide geometry is parametrized by cubic splines with four control parameters and allows for greater variability of the waveguides and corresponding dispersion relations. Moreover, it is more common that the third bandgap around $\kappa = 3\pi$ is opened. For discovering the formulae of either the dispersion relation or bandgap widths and centers, SR was employed. The discovered formulae significantly contribute to the overall readability of the system features and allow for use in design and optimization.

Chapter 5

Practical examples

To validate the discovered formulae and demonstrate some of their practical aspects, the results are compared with simulations performed using the Finite Element Method (FEM). First, the simulation setup is described in section 5.1. section 5.2 is dedicated to study how many unit cells are needed to reach performance close to one of an infinitely periodic structures. In section 5.3, the verification of discovered formulae is presented. Last but not least, illustrative examples of possible applications are shown in section 5.4.

5.1 Simulation setup

To validate the discovered formulae and demonstrate some of their practical aspects, the results are compared with simulations performed using FEM in COMSOL Multiphysics 5.5. For this purpose, the Helmholtz equation for acoustic pressure is simulated in a two-dimensional axially symmetric geometry. To allow the first three bandgaps to exist securely well below the cut-off frequency, the unit cell aspect ratio is chosen as $R_{\mathrm{max}}/\ell = 0.15$ unless stated otherwise. At the input to the system, a plane wave is prescribed using a background pressure field. The reflections are handled by perfectly matched layers at both ends of the domain. The amplitude transmission coefficient T is calculated using the magnitude of the complex sound pressure amplitude beyond the locally periodic structure.

5.2 Influence of the number of unit cells

As mentioned at the very beginning of the thesis in section 1.2, the perfect bandgap is obtained for an infinite structure. Hence, it will be first investigated how many unit cells are needed for complete attenuation, i.e., when the structure is sufficiently large that there is no transmission through the structure.

The geometry was chosen from the case in Figure 4.9 (middle row) providing a good fit. The simulation was not extended to other geometries due to its substantial computational cost. The transmission characteristics were calculated for locally periodic structures with different number of unit cells, namely of # unit cells, where $\# = \{5,\ 10, \ldots 50,\ 55\}$. Using higher number of unit cells is way more computationally demanding and the results prove that it would not provide any further value.

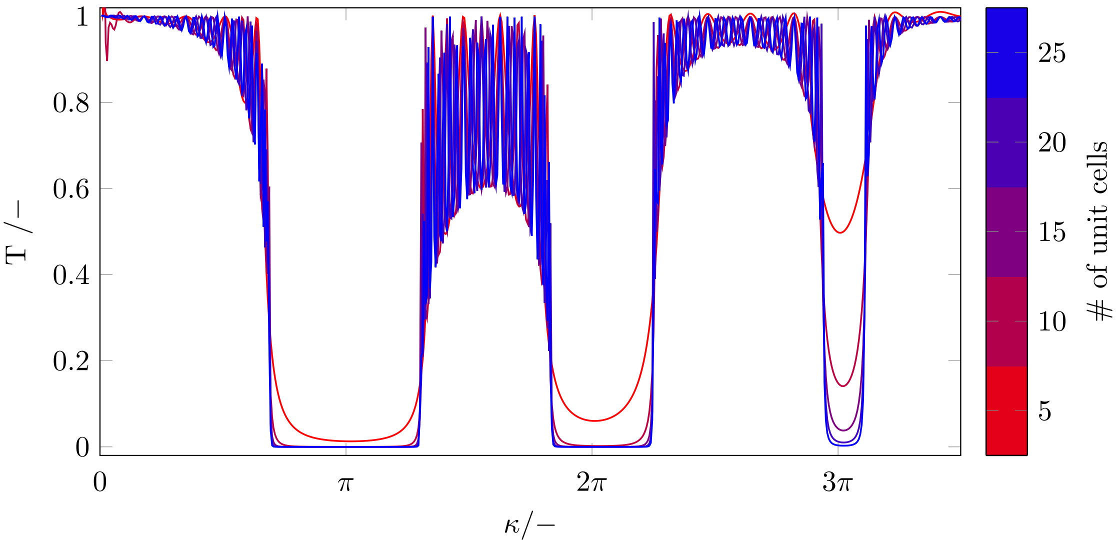


Figure 5.1: Transmission characteristics obtained from FEM simulations or various number of unit cells.

# cells	5	10	15	20	25	30	35	40	45	50	55
w_1 / Hz	-	93	100	103	103	103	103	103	103	104	104
w_2 / Hz	-	50	64	67	69	69	69	69	69	69	69
w_3 / Hz	-	-	-	-	18	22	24	25	26	26	26

Table 5.1: Bandgap widths for various number of unit cells, FEM simulations.

Figure 5.1 depicts the results of the simulations. The transmission characteristics of more than 25 unit cells are overlapping, and therefore are excluded from the image. Let the bandgap width be defined as the frequency range in Hz, where T $\leq 10^{-3}$ for the FEM simulation. Then the Table 5.1 contains the extracted bandgap widths.

Both the figure and the table show that the third bandgap is affected the most: it is necessary to have no less than 20 unit cells to obtain the transmission below 10^{-3}. However, to obtain consistently converging results, the amount of unit cells has to be at least twice as large. The second bandgap remains consistent from 25 unit cells on, the first from 20 on. For only 5 unit cells, none of the bandgaps reaches transmission below 10^{-3}. These results are consistent with the information from section 1.2

The Table 5.1 indicates that it would be necessary to perform the FEM simulations with at least 50 unit cells. However, due to the high computational demands of the FEM simulation, it was decided to limit the number of unit cells to 30 for the simulations presented in section 5.4, unless stated otherwise. One only has to keep in mind the downside of this approach: the third bandgap width will be slightly underestimated. An alternative approach is to simulate infinitely periodic structure by implementing only one unit cell with boundary conditions of Floquet-Bloch type, but then the practical insight would be lost.

5.3 Verification of discovered formulae

The discovered formulae are verified employing the results of FEM simulations of the Helmholtz equation for 50 unit cells and the geometry chosen from the case in Figure 4.9, middle row. Next, the bandgap widths are extracted from the dispersion relation predicted by the discovered formulae and also from the dataset obtained via Webster equation for unit cell with boundary conditions of Floquet-Bloch type.

In section 4.6, the bandgap widths w_1, w_2, w_3 were retrieved from the $\cos\xi$ as the width of range where $\cos\xi \leq -1$ for the first and the third gap and $\cos\xi \geq 1$ for the second gap. These conditions correspond to extracting the bandgap widths from dispersion relation as the frequency ranges where the imaginary part is nonzero.

Table 5.2 provides the comparison of the obtained bandgap widths. The first gap is underestimated by 1 Hz, which is probably an artefact of numerical precision. The second and third bandgap widths are overestimated by the Webster equation model compared to the FEM simulation. As noted in the previous section, this discrepancy is expected, since the Webster model assumes an infinite structure. The differences between the discovered formulae and the Webster equation model are dealt with in section 4.5.

	discovered formulae	Webster + FB	FEM simulation
w_1 / Hz	103	103	104
w_2 / Hz	82	73	69
w_3 / Hz	29	33	26

Table 5.2: Comparison of bandgap widths predicted by discovered formulae, the ground thruth from the dataset obtained via Webster equation for unit cell with boundary conditions of Floquet-Bloch type, and FEM simulations of the Helmholtz equation for 50 unit cells.

5.4 Examples of application

This section is dedicated to practical examples showing the power of discovered formulae.

5.4.1 Targeted bandgap with simple geometry constraints

First, a simple example is given of the application of Eqs. (4.15) and (4.21) to a made-up practice-motivated problem. Consider a circular ventilation pipe with a length of 8.5 meters and a diameter of 300 mm. The task is to attenuate (or ideally, completely block) a frequency band of 50 Hz centered around 210 Hz, which could correspond, e.g., to the blade passage frequency of a fan. Due to construction constraints, nothing can be added to the pipe diameter; only internal inserts are allowed. Furthermore, assume that the air inside the pipe is relatively cooler, with a sound speed of $c_0 = 335$ m s^{-1}.

First, using Eq. (4.15), the non-dimensional wavenumber for the first gap center frequency is calculated. For simplicity, let $r_d = 0$, then

$$m_1 = \pi + 0.38 \max(0.17, x_d r_d) = \pi + 0.38 \cdot 0.17 = 3.2\ . \tag{5.1}$$

Since $\kappa = 2\pi f\ell/c_0$, the length of a unit cell can be determined as

$$\ell = \frac{c_0 m_1}{2\pi \bar{f}} = \frac{335 \text{ m s}^{-1} \cdot 3.2}{2\pi \cdot 210 \text{ Hz}} = 0.81 \text{ m} \ , \tag{5.2}$$

which means, that on the given length of 8.5 m, the locally periodic structure can consist of only 10 unit cells.

Analogously, the dimensionless width w_1 corresponding to the 50 Hz band should have

$$w_1 = 2\pi f\ell/c_0 = 2\pi \cdot 50 \text{ Hz} \cdot 0.81 \text{ m} \ /335 \text{ m s}^{-1} = 0.76 \ . \tag{5.3}$$

This value can then be used in Eq. (4.21) to design the specific geometry.

While there are an infinite number of ways to satisfy this requirement, certain constraints must be considered. Notably, we are only considering inserts inside the pipe. If the pipe were part of a low Mach number flow ventilation system, a smooth profile would be required for hydrodynamic reasons. Therefore, the following choice becomes a natural one:

$$x_m = 0 \ , \qquad x_d = 0.33 \ , \qquad r_m = -0.29 \ , \qquad r_d = 0 \ . \tag{5.4}$$

Finally, these values must be converted back into dimensional geometric quantities. The results are presented in Fig. 5.2. Clearly, both the bandwidth and its center are appropriately sized, and the transmission coefficient is sufficiently low, making the design successful even with just ten unit cells.

However, in the general case, one may encounter the following problem in a locally periodic structure that is not accounted for by Eqs. (4.21)–(4.23). The bandgap widths derived from these equations assume an infinite periodic structure and do not account for the possibility that an evanescent wave might propagate far enough through the local structure. This could lead to a transmission coefficient greater than zero. Therefore, it is logical tohave formulae for the complete dispersion relation as well.

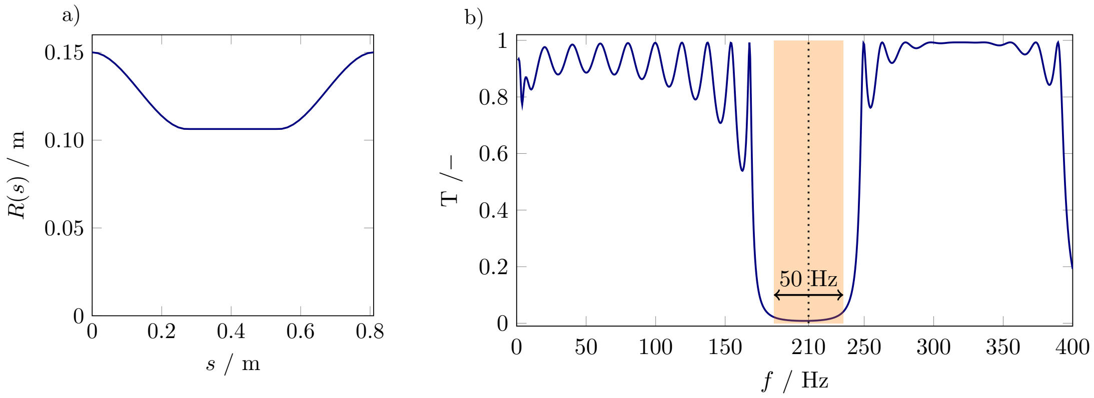


Figure 5.2: a) The radius function of the unit cell corresponding to the problem in Sec. 5.4.1; b) its transmission characteristic with the intended bandgap and its centre marked.

5.4.2 Three equally wide bandgaps

In the locally periodic structures studied, it is common for the width of the first bandgap to be larger than that of the second, and similarly for the third bandgap. To demonstrate the capabilities of the analytical formulae (4.21)–(4.23), a case with three bandgaps of approximately same width is shown.

To achieve this, we minimize the following objective function:

$$(w_1 - w_2)^2 + (w_1 - w_3)^2 + (w_2 - w_3)^2 + \max(0.5 - w_1, 0) \ , \tag{5.5}$$

using a simple coordinate descent optimization algorithm. The last term prevents the solution to be all the bandgap widths equal to zero. The resulting radius function and the transmission characteristic are depicted in Fig. 5.3.

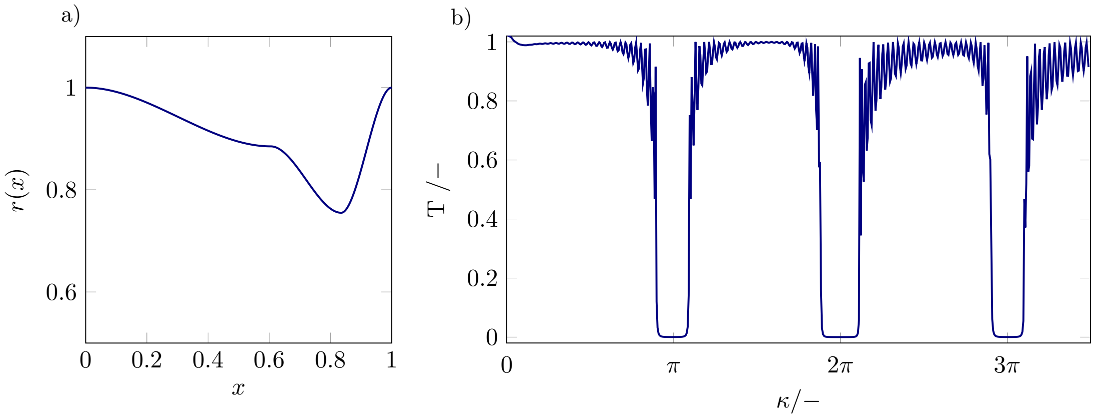


Figure 5.3: a) Radius function unit cell for the case of three equally wide bandgaps; b) the corresponding transmission characteristic.

5.4.3 Constant second bandgap with different geometries

Some special cases can be derived from the relatively complicated full form of Eqs. (4.21)–(4.23). Let's derive an example, where the second bandgap remains constant for different geometries, assuming a unit cell consists of a protrusion with a flat tip, i.e. $x_m = 0$, $r_d = 0$ and $r_m > 0$. For these assumptions, the total differential of Eq. (4.22) reads:

$$\mathrm{d}w_2 = (4.95r_m - 0.2)\ \mathrm{d}x_d + 4.95x_d\ \mathrm{d}r_m \ . \tag{5.6}$$

One possible application of this equation is to guide the adjustment of r_m and x_d in such a way that w_2 remains constant. This effect is depicted in Fig. 5.4, where the parameters of the second bandgap remain fixed, while those of the first and third bandgaps vary. The behavior of the third bandgap can be understood by noting that it (corresponding to relatively shorter wavelengths) is more effectively supported by localized changes in the radius function. In essence, the red radius curves in Fig. 5.4 feature a larger section of the waveguide with a constant radius, with steeper variations occurring at the beginning and end of the unit cell.

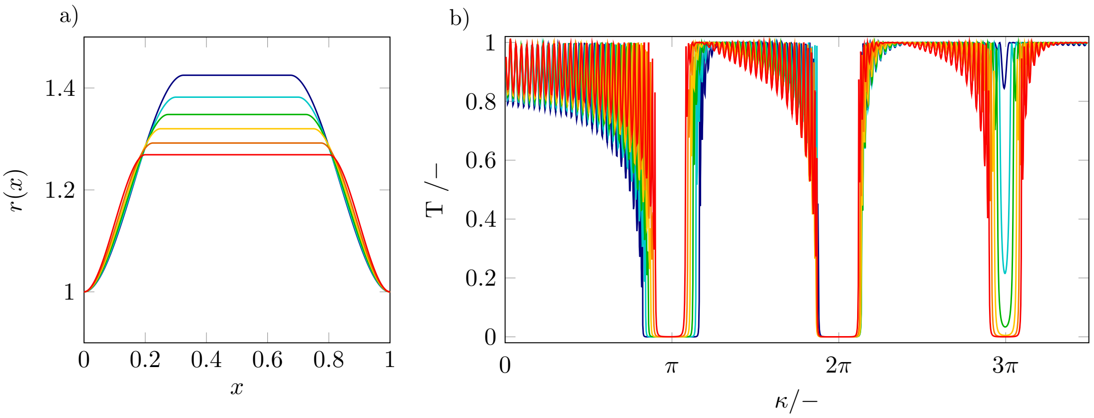


Figure 5.4: a) Different unit cell geometries designed to satisfy $\mathrm{d}w_2 = 0$ in Eq. (5.6). b) The corresponding transmission characteristics.

5.4.4 Related problems

It was mentioned in section 2.1, that an axisymmetric waveguide is considered throughout the work, i.e. it holds for the cross-section function $A(s) = \pi R(s)^2$, where $R(x)$ is the radius function and the s the spatial coordinate along the waveguide axis. However, the results are not limited to axisymmetric waveguides. The modification that transforms the cross-sectional area function $A(x)$ into the radius function $R(x)$

$$\frac{1}{A(x)}\frac{\mathrm{d}A(x)}{\mathrm{d}x} \rightarrow \frac{2}{R(x)}\frac{\mathrm{d}R(x)}{\mathrm{d}x} \tag{5.7}$$

holds for any cross-sectional shape described by a single linear parameter, such as the side of a square, a rectangle with a fixed aspect ratio, or an ellipse with constant eccentricity.

Chapter 6

Conclusions

The thesis dealt with modeling band structure properties in sonic crystals using machine learning, with the focus on one-dimensional sonic crystals of continuous geometries governed by the Webster equation. The aim of this thesis was to discover an analytical model of band structure in such structures to improve the overall readability of the system features and enable more efficient optimization. This was due to the fact that, up until now, the device design and optimization were possible only with computationally demanding numerical optimization repeatedly going back and forth from waveguide geometry to the dispersion relation, without much knowledge about the influence of the parameters on geometry.

Having an access to a numerical model provided by the Webster equation with periodic boundary conditions of Floquet-Bloch type is a great opportunity for data-driven discovery of underlying patterns. Based on the principle of parsimony (also known as the Ockham's razor principle), it can be assumed that the model should be sparse in terms of possible functions in the governing equations. Involving current knowledge of the problem's underlying physics and machine learning techniques such as coordinate transformations and sparse regression, the governing equations can be inferred from the dataset.

Initially, when parametrizing geometry by sine to the second power in chapter 3, the formulae were extracted from the dataset after coordinate transformation by PCA using LASSO and Levenberg-Marquardt algorithm in combination with a genetic algorithm CMA-ES. However, this parametrization did not provide enough variability of the geometry, which in turn meant that the control of bandgaps was not sufficient. Therefore in chapter 4, a more general parametrization of the waveguide geometry was employed, namely cubic splines with four control parameter. Due to the increased number of degrees of freedom, a more advanced method was necessary for extracting formulae from the dataset. For this task, the symbolic regression was chosen, providing a couple of mathematical expression sorted by complexity and accuracy. Based on the tradeoff between these two, the final model was selected.

The results were validated by employing FEM to solve the Helmholtz equation for the acoustic pressure in COMSOL Multiphysics 5.5 and shown to be supported by physical interpretation. It was verified, that it is possible to model the band structure properties in sonic crystals with smooth geometry from the dataset employing machine learning.

The results of the proof of concept from chapter 3 were presented at The 28th International Student Conference on Electrical Engineering at Faculty of Electrical Engineering, CTU Prague, and published in the conference proceedings [64]. In the Best Poster Award, the poster was awarded with first prize in the Natural Sciences section. The early stage results from chapter 4 were presented at The 53rd International Congress and Exposition on Noise Control Engineering in Nantes, France, and published in the conference proceedings [65]. The final results with examples of formulae application were published in a Q1-ranked journal, The Journal of Sound and Vibration [57].

One of the future goals is to design a custom loss function, that would penalize the formulae not supported by physical interpretation. In this way, the symbolic regression would provide us more possible solutions and it would not be necessary to select the physically viable by hand.

Moreover, it is worth considering transforming the problem to a Schrödinger-like equation and defining the geometry not by the radius function as in this thesis, but by the gaussian curvature. The preliminary results of this approach were presented at the 51st Annual Meeting on Acoustics DAS/DAGA 2025 in Copenhagen this March, and published in the conference proceedings [66]. Transforming the problem to the Schrödinger-like equation allows potentially for a more multiphysical approach, because it presents a common ground for various problems of propagation through inhomogeneous structures.

Appendix A

References

References to the author's work are listed first, followed by other references cited within this work.

A.1 Thesis core publications

Thesis-related articles in peer-reviewed journals

[57] Viktor Hruška, Aneta Furmanová, and Michal Bednařík. "Analytical formulae for design of one-dimensional sonic crystals with smooth geometry based on symbolic regression". In: *Journal of Sound and Vibration* 597 (Feb. 2025), p. 118821. ISSN: 0022-460X. DOI: `10.1016/j.jsv.2024.118821`.

Thesis-related conference articles

[64] A. Furmanová. "Acoustic transmission modelling in locally periodic structures employing machine learning". In: *Proceedings of the International Student Scientific Conference Poster - 28/2024. 28th International Student Conference on Electrical Engineering. POSTER 2024.* Prague: ČVUT. Fakulta elektrotechnická, May 23, 2024. ISBN: 978-80-01-07299-8.

[65] Aneta Furmanová, Viktor Hruška, and Michal Bednařík. "Machine learning for locally periodic structure transmission modelling". In: *INTER-NOISE and NOISE-CON Congress and Conference Proceedings.* Vol. 270. Institute of Noise Control Engineering, Oct. 2024, pp. 3099–3105. DOI: `10.3397/IN_2024_3274`.

[66] Aneta Furmanová and Viktor Hruška. "Modelling Band Structure Properties Using Physics-Informed Machine Learning". In: *Proceedings of DAS/DAGA 2025.* Berlin: Deutsche Gesellschaft für Akustik e.V. (DEGA), 2025, pp. 998–1000. ISBN: 978-3-939296-23-2. DOI: `10.71568/DASDAGA2025.605`.

A.2 Cited references

[1] Sajeev John. “Strong localization of photons in certain disordered dielectric superlattices”. In: *Phys. Rev. Lett.* 58 (23 June 1987), pp. 2486–2489. DOI: `10.1103/PhysRevLett.58.2486`.

[2] Eli Yablonovitch. “Inhibited Spontaneous Emission in Solid-State Physics and Electronics”. In: *Phys. Rev. Lett.* 58 (20 May 1987), pp. 2059–2062. DOI: `10.1103/PhysRevLett.58.2059`.

[3] José Sánchez-Dehesa and Arkadii Krokhin. “Introduction to Acoustics of Phononic Crystals. Homogenization at Low Frequencies”. In: *Phononic Crystals: Fundamentals and Applications*. Ed. by Abdelkrim Khelif and Ali Adibi. New York, NY: Springer New York, 2016, pp. 1–21. ISBN: 978-1-4614-9393-8. DOI: `10.1007/978-1-4614-9393-8_1`.

[4] M.M. Sigalas and E.N. Economou. “Elastic and acoustic wave band structure”. In: *Journal of Sound and Vibration* 158.2 (1992), pp. 377–382. ISSN: 0022-460X. DOI: `10.1016/0022-460X(92)90059-7`.

[5] M. S. Kushwaha et al. “Acoustic band structure of periodic elastic composites”. In: *Phys. Rev. Lett.* 71 (13 Sept. 1993), pp. 2022–2025. DOI: `10.1103/PhysRevLett.71.2022`.

[6] R. Martínez-Sala et al. “Sound attenuation by sculpture”. In: *Nature* 378.6554 (Nov. 1995), pp. 241–241. ISSN: 1476-4687. DOI: `10.1038/378241a0`. URL: <`http://dx.doi.org/10.1038/378241a0`>.

[7] Arpan Gupta. “A review on sonic crystal, its applications and numerical analysis techniques”. In: *Acoustical Physics* 60.2 (Mar. 2014), pp. 223–234. ISSN: 1562-6865. DOI: `10.1134/s1063771014020080`.

[8] Jiangyi Zhang, Bo Hu, and Shibo Wang. “Review and perspective on acoustic metamaterials: From fundamentals to applications”. In: *Applied Physics Letters* 123.1 (July 2023). ISSN: 1077-3118. DOI: `10.1063/5.0152099`.

[9] Muhammad and C. W. Lim. “From Photonic Crystals to Seismic Metamaterials: A Review via Phononic Crystals and Acoustic Metamaterials”. In: *Archives of Computational Methods in Engineering* 29.2 (June 2021), pp. 1137–1198. ISSN: 1886-1784. DOI: `10.1007/s11831-021-09612-8`.

[10] Carlos III University of Madrid. *A new type of acoustic insulation enables sound to be concentrated in corners.* Phys.org, October 29, 2019. Oct. 2019. URL: <`https://phys.org/news/2019-10-acoustic-insulation-enables-corners.html`>. Accessed 22-Apr-2025.

[11] V. Romero-Garcia and A.C. Hladky-Hennion. *Fundamentals and Applications of Acoustic Metamaterials: From Seismic to Radio Frequency.* Wiley, 2019. ISBN: 9781786303363.

[12] Vicent Romero-García. “Sound Wave Propagation in Sonic Crystals”. In: *Acoustic Waves in Periodic Structures, Metamaterials, and Porous Media: From Fundamentals to Industrial Applications.* Ed. by Noé Jiménez, Olga Umnova, and Jean-Philippe Groby. Cham: Springer International Publishing, 2021, pp. 65–102. ISBN: 978-3-030-84300-7. DOI: `10.1007/978-3-030-84300-7_3`.

[13] R. Lucklum and J. Li. “Phononic crystals for liquid sensor applications”. In: *Measurement Science and Technology* 20.12 (2009). ISSN: 09570233. DOI: `10.1088/0957-0233/20/12/124014`.

[14] R. Lucklum, J. Li, and M. Zubtsov. “1D and 2D phononic crystal sensors”. In: *Procedia Engineering* 5 (2010), pp. 436–439. ISSN: 18777058. DOI: `10.1016/j.proeng.2010.09.140`.

[15] A. Gueddida et al. “Tubular phononic crystal sensor”. In: *Journal of Applied Physics* 130.10 (2021). ISSN: 10897550. DOI: `10.1063/5.0051660`.

[16] Dipen Sinha and Cristian Pantea. “Broadband directional ultrasound propagation using sonic crystal and nonlinear medium”. In: *Proceedings of Meetings on Acoustics* 19 (2013). ISSN: 1939800X. DOI: `10.1121/1.4799990`.

[17] J. Bonhomme et al. “Love waves dispersion by phononic pillars for nano-particle mass sensing”. In: *Applied Physics Letters* 114.1 (Jan. 2019). ISSN: 1077-3118. DOI: `10.1063/1.5068681`.

[18] Ralf Lucklum, Mikhail Zubtsov, and Aleksandr Oseev. “Phoxonic crystals—a new platform for chemical and biochemical sensors”. In: *Analytical and Bioanalytical Chemistry* 405.20 (June 2013), pp. 6497–6509. ISSN: 1618-2650. DOI: `10.1007/s00216-013-7093-9`.

[19] Fuyin Ma et al. “Acoustic focusing and imaging via phononic crystal and acoustic metamaterials”. In: *Journal of Applied Physics* 131.1 (2022). ISSN: 10897550. DOI: `10.1063/5.0074503`.

[20] A. Sukhovich et al. “Experimental and Theoretical Evidence for Subwavelength Imaging in Phononic Crystals”. In: *Physical Review Letters* 102 (15 Apr. 2009), p. 154301. DOI: `10.1103/PhysRevLett.102.154301`.

[21] Hrishikesh Danawe and Serife Tol. “Broadband subwavelength imaging of flexural elastic waves in flat phononic crystal lenses”. In: *Scientific Reports* 13.1 (May 2023). ISSN: 2045-2322. DOI: `10.1038/s41598-023-34314-5`.

[22] Yuqi Jin and Arup Neogi. “Ultrasound Imaging by Thermally Tunable Phononic Crystal Lens”. In: *International Journal of Molecular Sciences* 22.15 (July 2021), p. 7966. ISSN: 1422-0067. DOI: `10.3390/ijms22157966`.

[23] Benjamin Bühling, Stefan Maack, and Christoph Strangfeld. “Using sonic crystals to separate the acoustic from the flow field of a fluidic transducer”. In: *Applied Acoustics* 189 (2022). ISSN: 1872910X. DOI: `10.1016/j.apacoust.2021.108608`.

[24] M. Thota and K. W. Wang. “Reconfigurable origami sonic barriers with tunable bandgaps for traffic noise mitigation”. In: *Journal of Applied Physics* 122.15 (Oct. 2017). ISSN: 1089-7550. DOI: `10.1063/1.4991026`.

[25] Rosa Martínez-Sala et al. “Control of noise by trees arranged like sonic crystals”. In: *Journal of Sound and Vibration* 291.1 (2006), pp. 100–106. ISSN: 0022-460X. DOI: `10.1016/j.jsv.2005.05.030`.

[26] Youzhi Jiang et al. “Vibration attenuation analysis of periodic underground barriers using complex band diagrams”. In: *Computers and Geotechnics* 128 (2020), p. 103821. ISSN: 0266-352X. DOI: `10.1016/j.compgeo.2020.103821`.

[27] NK Batra, P Matic, and RK Everett. “Sonic crystal composites for selective noise reduction”. In: *2002 IEEE Ultrasonics Symposium, 2002. Proceedings.* Vol. 1. IEEE. 2002, pp. 547–550.

[28] MS Kushwaha et al. “Sonic stop-bands for cubic arrays of rigid inclusions in air”. In: *The European Physical Journal B-Condensed Matter and Complex Systems* 3.2 (1998), pp. 155–161.

[29] María Peiro-Torres et al. “Open Noise Barriers Based on Sonic Crystals. Advances in Noise Control in Transport Infrastructures”. In: *Transportation Research Procedia* 18 (Dec. 2016), pp. 392–398. DOI: `10.1016/j.trpro.2016.12.051`.

[30] Carlos Albino et al. “3D FEM analysis of the effect of buried phononic crystal barriers on vibration mitigation”. In: *Engineering Structures* 196 (2019), p. 109340. ISSN: 0141-0296. DOI: `10.1016/j.engstruct.2019.109340`.

[31] O. Casablanca. “Periodic foundation piles for the seismic protection of structures”. In: *Soil Dynamics and Earthquake Engineering* 182 (2024), p. 108742. ISSN: 0267-7261. DOI: `10.1016/j.soildyn.2024.108742`.

[32] A. Gupta, K. M. Lim, and C. H. Chew. “Analysis of frequency band structure in one-dimensional sonic crystal using Webster horn equation”. In: *Applied Physics Letters* 98.20 (2011). ISSN: 1077-3118. DOI: `10.1063/1.3592570`.

[33] Muhammad, John Kennedy, and C.W. Lim. “Machine learning and deep learning in phononic crystals and metamaterials – A review”. In: *Materials Today Communications* 33 (Dec. 2022), p. 104606. ISSN: 2352-4928. DOI: `10.1016/j.mtcomm.2022.104606`.

[34] Bin Zheng et al. “Inverse design of acoustic metamaterials based on machine learning using a Gauss–Bayesian model”. In: *Journal of Applied Physics* 128.13 (Oct. 2020). ISSN: 1089-7550. DOI: `10.1063/5.0012392`.

[35] Johannes D. Schmid, Jakob Wagner, and Steffen Marburg. “Neural Operators for the Design of Acoustic Metamaterials”. In: *Proceedings of DAS|DAGA 2025*. Berlin: Deutsche Gesellschaft für Akustik e.V. (DEGA), 2025, pp. 585–587. ISBN: 978-3-939296-23-2. DOI: `10.71568/DASDAGA2025.280`.

[36] P. D. C. King and T. J. Cox. “Acoustic band gaps in periodically and quasiperiodically modulated waveguides”. In: *Journal of Applied Physics* 102.1 (July 2007). ISSN: 1089-7550. DOI: `10.1063/1.2749483`.

[37] Vincent Laude. *Phononic Crystals: Artificial Crystals for Sonic, Acoustic, and Elastic Waves.* De Gruyter, May 2020. ISBN: 9783110641189. DOI: `10.1515/9783110641189`.

[38] Mihail Sigalas et al. “Classical vibrational modes in phononic lattices: theory and experiment”. In: *Zeitschrift für Kristallographie - Crystalline Materials* 220.9–10 (Oct. 2005), pp. 765–809. ISSN: 2194-4946. DOI: `10.1524/zkri.2005.220.9-10.765`.

[39] A.R. Diaz, A.G. Haddow, and L. Ma. “Design of band-gap grid structures”. In: *Structural and Multidisciplinary Optimization* 29.6 (Jan. 2005), pp. 418–431. ISSN: 1615-1488. DOI: `10.1007/s00158-004-0497-6`.

[40] Ole Sigmund and Jakob Jensen. "Systematic design of phononic band-gap materials and structures by topology optimization". In: *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 361 (June 2003), pp. 1001–19. DOI: `10.1098/rsta.2003.1177`.

[41] George A. Gazonas et al. "Genetic algorithm optimization of phononic bandgap structures". In: *International Journal of Solids and Structures* 43.18 (2006), pp. 5851–5866. ISSN: 0020-7683. DOI: `10.1016/j.ijsolstr.2005.12.002`.

[42] Steven L. Brunton and J. Nathan Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control.* 2nd ed. Cambridge University Press, 2022.

[43] George Em Karniadakis et al. "Physics-informed machine learning". In: *Nature Reviews Physics* 3.6 (May 2021), pp. 422–440. ISSN: 2522-5820. DOI: `10.1038/s42254-021-00314-5`.

[44] Gustau Camps-Valls et al. "Discovering causal relations and equations from data". In: *Physics Reports* 1044 (2023), pp. 1–68. ISSN: 0370-1573. DOI: `10.1016/j.physrep.2023.10.005`.

[45] Silviu-Marian Udrescu and Max Tegmark. "AI Feynman: A physics-inspired method for symbolic regression". In: *Science Advances* 6.16 (2020), eaay2631. DOI: `10.1126/sciadv.aay2631`.

[46] Michael Schmidt and Hod Lipson. "Distilling free-form natural laws from experimental data". In: *Science* 324.5923 (2009), pp. 81–85. DOI: `10.1126/science.1165893`.

[47] David T. Blackstock. *Fundamentals of physical acoustics.* Wiley, 2000, p. 422. ISBN: 978-0-471-31979-5.

[48] G. Floquet. "Sur les équations différentielles linéaires à coefficients périodiques". In: *Annales scientifiques de l'École normale supérieure* 12 (1883), pp. 47–88. ISSN: 1873-2151. DOI: `10.24033/asens.220`.

[49] Felix Bloch. "Über die Quantenmechanik der Elektronen in Kristallgittern". In: *Zeitschrift für Physik* 52.7–8 (July 1929), pp. 555–600. ISSN: 1434-601X. DOI: `10.1007/bf01339455`.

[50] Sjoerd W. Rienstra. "Webster's Horn Equation Revisited". In: *SIAM Journal on Applied Mathematics* 65.6 (Jan. 2005), pp. 1981–2004. ISSN: 1095-712X. DOI: `10.1137/s0036139902413040`.

[51] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: `10.1038/s41586-020-2649-2`.

[52] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: `10.1038/s41592-019-0686-2`.

[53] M. Bednarik and M. Cervenka. "A wide class of analytical solutions of the Webster equation". In: *Journal of Sound and Vibration* 469 (2020), p. 115169. ISSN: 0022-460X. DOI: `10.1016/j.jsv.2019.115169`.

[54] Viktor Hruska, Michal Reiser, and Michal Bednarik. "Fast design of non-uniform acoustic waveguides with prescribed eigenfrequencies". In: *Wave Motion* 122 (Oct. 2023), p. 103192. ISSN: 0165-2125. DOI: `10.1016/j.wavemoti.2023.103192`.

[55] Nikolaus Hansen and Andreas Ostermeier. "Completely Derandomized Self-Adaptation in Evolution Strategies". In: *Evolutionary Computation* 9.2 (2001), pp. 159–195. DOI: `10.1162/106365601750190398`.

[56] Robert Tibshirani. "Regression Shrinkage and Selection Via the Lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (Dec. 2018), pp. 267–288. ISSN: 0035-9246. DOI: `10.1111/j.2517-6161.1996.tb02080.x`.

[58] Gabriel Kronberger et al. *Symbolic Regression*. July 2024. ISBN: 9781315166407. DOI: `10.1201/9781315166407`.

[59] Marco Virgolin and Solon P Pissis. "Symbolic Regression is NP-hard". In: *Transactions on Machine Learning Research* (2022). ISSN: 2835-8856.

[60] John R. Koza. "Genetic programming as a means for programming computers by natural selection". In: *Statistics and Computing* 4.2 (June 1994), pp. 87–112. ISSN: 0960-3174. DOI: `10.1007/BF00175355`.

[61] Jan Žegklitz and Petr Pošík. "Benchmarking state-of-the-art symbolic regression algorithms". In: *Genetic Programming and Evolvable Machines* 22.1 (Mar. 2020), pp. 5–33. ISSN: 1573-7632. DOI: `10.1007/s10710-020-09387-0`.

[62] Guilherme S. Imai Aldeia et al. "Call for Action: towards the next generation of symbolic regression benchmark". In: (2025). DOI: `10.48550/ARXIV.2505.03977`.

[63] Miles Cranmer. *Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl*. 2023. DOI: `10.48550/ARXIV.2305.01582`.